## FIGURE 1

CGGACGCGTGGGTGCGAGGCGAAGGTGACCGGGGACCGAGCATTTCAGATCTGCTCGGTAGA CCTGGTGCACCACCACCATGTTGGCTGCAAGGCTGGTGTCTCCGGACACTACCTTCTAGG GTTTTCCACCCAGCTTTCACCAAGGCCTCCCCTGTTGTGAAGAATTCCATCACGAAGAATCA ATGGCTGTTAACACCTAGCAGGGAATATGCCACCAAAACAAGAATTGGGATCCGGCGTGGGA GAACTGGCCAAGAACTCAAAGAGGCAGCATTGGAACCATCGATGGAAAAAATATTTAAAATT GATCAGATGGGAAGATGGTTTGTTGCTGGAGGGGCTGCTGTTGGTCTTGGAGCATTGTGCTA CTATGGCTTGGGACTGTCTAATGAGATTGGAGCTATTGAAAAGGCTGTAATTTGGCCTCAGT ATGTCAAGGATAGAATTCATTCCACCTATATGTACTTAGCAGGGAGTATTGGTTTAACAGCT TTGTCTGCCATAGCAATCAGCAGAACGCCTGTTCTCATGAACTTCATGATGAGAGGCTCTTG GGTGACAATTGGTGTGACCTTTGCAGCCATGGTTGGAGCTGGAATGCTGGTACGATCAATAC CATATGACCAGAGCCCAGGCCCAAAGCATCTTGCTTGGTTGCTACATTCTGGTGTGATGGGT GCAGTGGTGGCTCCTCTGACAATATTAGGGGGTCCTCTTCTCATCAGAGCTGCATGGTACAC AGCTGGCATTGTGGGAGGCCTCTCCACTGTGGCCATGTGTGCGCCCAGTGAAAAGTTTCTGA ACATGGGTGCACCCCTGGGAGTGGGCCTGGGTCTCGTCTTTGTGTCCTCATTGGGATCTATG TTTCTTCCACCTACCACCGTGGCTGGTGCCACTCTTTACTCAGTGGCAATGTACGGTGGATT AGTTCTTTTCAGCATGTTCCTTCTGTATGATACCCAGAAAGTAATCAAGCGTGCAGAAGTAT CACCAATGTATGGAGTTCAAAAATATGATCCCATTAACTCGATGCTGAGTATCTACATGGAT AAGTGACTCAGCTTCTGGCTTCTCTGCTACATCAAATATCTTGTTTAATGGGGCAGATATGC ATTAAATAGTTTGTACAAGCAGCTTTCGTTGAAGTTTAGAAGATAAGAAACATGTCATCATA TTTAAATGTTCCGGTAATGTGATGCCTCAGGTCTGCCTTTTTTTCTGGAGAATAAATGCAGT AATCCTCTCCCAAATAAGCACACACATTTTCAATTCTCATGTTTGAGTGATTTTAAAATGTT TTGGTGAATGTGAAAACTAAAGTTTGTGTCATGAGAATGTAAGTCTTTTTTCTACTTTAAAA GAGTCACCTGCAGTCTTTTGTTTTTTTAAATACTTAGAACTTAGCACTTGTGTTATTGATTA GCTGAACTTAACAAACTGTTCATCCTGAAACAGGCACAGGTGATGCATTCTCCTGCTGTTG CTTCTCAGTGCTCTCTTTCCAATATAGATGTGGTCATGTTTGACTTGTACAGAATGTTAATC ATACAGAGAATCCTTGATGGAATTATATATGTGTGTTTTACTTTTGAATGTTACAAAAGGAA ATAACTTTAAAACTATTCTCAAGAGAAAATATTCAAAGCATGAAATATGTTGCTTTTTCCAG AATACAAACAGTATACTCATG

# FIGURE 2

MLAARLVCLRTLPSRVFHPAFTKASPVVKNSITKNQWLLTPSREYATKTRIGIRRGRTGQEL
KEAALEPSMEKIFKIDQMGRWFVAGGAAVGLGALCYYGLGLSNEIGAIEKAVIWPQYVKDRI
HSTYMYLAGSIGLTALSAIAISRTPVLMNFMMRGSWVTIGVTFAAMVGAGMLVRSIPYDQSP
GPKHLAWLLHSGVMGAVVAPLTILGGPLLIRAAWYTAGIVGGLSTVAMCAPSEKFLNMGAPL
GVGLGLVFVSSLGSMFLPPTTVAGATLYSVAMYGGLVLFSMFLLYDTQKVIKRAEVSPMYGV
QKYDPINSMLSIYMDTLNIFMRVATMLATGGNRKK

### FIGURE 3

GAAGGCTGCCTCGCTGGTCCGAATTCGGTGGCGCCACGTCCGCCCGTCTCCGCCTTCTGCAT GGTCGGCACGGGGGGTCTTGTGCATCTTGGCTACCTGTGGGTCGAAGATGTCGG ACATCGGAGACTGGTTCAGGAGCATCCCGGCGATCACGCGCTATTGGTTCGCCGCCACCGTCGCCGTGCCCTTGGTCGGCAAACTCGGCCTCATCAGCCCGGCCTACCTCTTCCTCTGGCCCGA AGCCTTCCTTTATCGCTTTCAGATTTGGAGGCCAATCACTGCCACCTTTTATTTCCCTGTGG TGCTGATCAGAATGGCGGAGGCGGGAGACACAACTGGGGCCCAGGGCTTTCGACTTGGAGACC ACAACAATCATATTCACGTTATTTTCCCCTTTTGGTGGCAGAACTGTTACCAATAGGGGGAG AAGACAGCCACGGATGAAGCGTTTCTCAGCTTTTGGAATTGCTTCGACTGACATCCGTTGTT AGGGAATAACATGATTTAAGGTTGAAATGGCTTTAGAATCATTTGGGTTTGAGGGTGTGTTA CAGTCATTTTTTCCTAAAGGTTTACAAGTATTTAGAACTTTTCAGTTCAGGGCAAAATGTTC ATGAAGTTATTCCTCTTAAACATGGTTAGGAAGCTGATGACGTTATTGATTTTGTCTGGATT ATGTTTCTGGAATAATTTTACCAAAACAAGCTATTTGAGTTTTGACTTGACAAGGCAAAACA TGACAGTGGATTCTCTTTACAAATGGAAAAAAAAAATCCTTATTTTGTATAAAGGACTTCCC TTTTTGTAAACTAATCCTTTTTATTGGTAAAAATTGTAAAATTAAAATGTGCAACTTG

# FIGURE 4

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGLISPAYLFLWPEAFLYRFQIWRPITATFYF PVGPGTGFLYLVNLYFLYQYSTRLETGAFDGRPADYLFMLLFNWICIVITGLAMDMQLLMIP LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIGGSVINELIGNLVGHLYFFL MFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRL GDQ

#### Transmembrane domain:

amino acids 98-116, 152-172

### N-myristoylation site.

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

### Glycosaminoglycan attachment site.

amino acids 218-222

## FIGURE 5

GGGGCCGCGGTCTAGGGCGGCTACGTGTTTTGCCATAGCGACCATTTTTGCATTAACTGGTTG GTAGCTTCTATCCTGGGGGCTGAGCGACTGCGGGCCAGCTCTTCCCCTACTCCCTCTCGGCT CCTTGTGGCCCAAAGGCCTAACCGGGGTCCGGCGGTCTGGCCTAGGGATCTTCCCCGTTGCC CCTTTGGGGCGGG ATG GCTGCGGAAGAAGAAGACGAGGTGGAGTGGGTAGTGGAGAGCATCGCGGGGTTCCTGCGAGGCCCAGACTGGTCCATCCCCATCTTGGACTTTGTGGAACAGAAATGT GAAGTTAACTGCAAAGGAGGGCATGTGATAACTCCAGGAAGCCCAGAGCCGGTGATTTTGGT GGCCTGTGTTCCCCCTTGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGATTC ATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAATT AATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAGGC CATTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCCAGA AAAACATTGAAATGCAGCTGCAAGCCATTCGAATAATTCAAGAGAGAAATGGTGTATTACCT GACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAATCCT GAGGGAAGTTCTTAGAAAATCAAAAGAGGAATATGACCAGGAAGAAGAAAGGAAGAAGAAGAAAA AACAGTTATCAGAGGCTAAAACAGAAGAGCCCACAGTGCATTCCAGTGAAGCTGCAATAATG CACAAAAAGGCCTGAAGATTCCTGGCTTAGAGCATGCGAGCATTGAAGGACCAATAGCAAAC TTATCAGTACTTGGAACAGAAGAACTTCGGCAACGAGAACACTATCTCAAGCAGAAGAGAGA TAAGTTGATGTCCATGAGAAAGGATATGAGGACTAAACAGATACAAAATATGGAGCAGAAAG GAAAACCCACTGGGGAGGTAGAGGAAATGACAGAGAAACCAGAAATGACAGCAGAGGAGAAAG CAAACATTACTAAAGAGGAGATTGCTTGCAGAGAAACTCAAAGAAGAAGTTATTAATAAG**TA** CTTACACTG

## FIGURE 6

MAAEEEDEVEWVVESIAGFLRGPDWSIPILDFVEQKCEVNCKGGHVITPGSPEPVILVACVP LVFDDEEESKLTYTEIHQEYKELVEKLLEGYLKEIGINEDQFQEACTSPLAKTHTSQAILQP VLAAEDFTIFKAMMVQKNIEMQLQAIRIIQERNGVLPDCLTDGSDVVSDLEHEEMKILREVL RKSKEEYDQEEERKRKKQLSEAKTEEPTVHSSEAAIMNNSQGDGEHFAHPPSEVKMHFANQS IEPLGRKVERSETSSLPQKGLKIPGLEHASIEGPIANLSVLGTEELRQREHYLKQKRDKLMS MRKDMRTKQIQNMEQKGKPTGEVEEMTEKPEMTAEEKQTLLKRRLLAEKLKEEVINK

### N-glycosylation sites.

amino acids 224-228, 246-250, 285-289

### N-myristoylation site.

amino acids 273-279

#### Amidation site.

amino acids 252-256

## Cytosolic fatty-acid binding proteins.

amino acids 78-108

## FIGURE 7

GGGCACAGCACATGTGAAGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGAT
TCATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAA
TTAATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAG
GCCATTTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCC
AGAAAAACATTGAAATGCAGCTGCAAGCCATTCGAATAATTCAAGAGAGAAAATGGTGTATTA
CCTGACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAAT
CCTGAGGGAAGTTCTTAGAAAAATCAAAAGAGGAATATGACCAGGAA

## FIGURE 8

TAGCTTCTCCACGTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTG AGTGGAATGGAAAAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATAC ATTCCTGCTAGGTGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATT CTGCCAATGAAGAAAACAAGTATGATTATCTTCCAACTACTGTGAATGTGTGCTCAĢAACTG AAATTTGAAATATGCTTCCTGGAAGGAATTCTCTGATTTCATGAAGTGGTCCATTCCTGCCT TTCTTTATTTCCTGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATG GCTGTTATCTTCTCAAATTTTAGCATTATAACAACAGCTCTTCTATTCAGGATAGTGCTGAA TGACTGCCGGGACTAAAACTTTACAGCACAACTTGGCAGGACGTGGATTTCATCACGATGCC TACAGCAAAGGAATGGACTTTTCCTGAAGCTAAATGGAACACCACAGCCAGAGTTTTCAGTC ACATCCGTCTTGGCATGGGCCATGTTCTTATTATAGTCCAGTGTTTTATTTCTTCAATGGCT AATATCTATAATGAAAAGATACTGAAGGAGGGGAACCAGCTCACTGAAAGCATCTTCATACA GAACAGCAAACTCTATTTCTTTGGCATTCTGTTTAATGGGCTGACTCTGGGCCTTCAGAGGA CTTATTTTTGTAACTGCATTCCAGGGCCTTTCAGTGGCTTTCATTCTGAAGTTCCTGGATAA TCTTTGACTTCAGGCCCTCCCTGGAATTTTTCTTGGAAGCCCCATCAGTCCTTCTCTCTATA TTTATTTATAATGCCAGCAAGCCTCAAGTTCCGGAATACGCACCTAGGCAAGAAAGGATCCG AGATCTAAGTGGCAATCTTTGGGAGCGTTCCAGTGGGGATGGAGAAGAACTAGAAAGACTTA  $\tt CCAAACCCAAGAGTGATGAGTCAGATGAAGATACTTTC{\color{red}{\textbf{TAA}}} CTGGTACCCACATAGTTTGCA$ GCTCTCTTGAACCTTATTTTCACATTTTCAGTGTTTTGTAATATTTATCTTTTCACTTTGATA AACCAGAAATGTTTCTAAATCCTAATATTCTTTGCATATATCTAGCTACTCCCTAAATGGTT CCATCCAAGGCTTAGAGTACCCAAAGGCTAAGAAATTCTAAAGAACTGATACAGGAGTAACA ATATGAAGAATTCATTAATATCTCAGTACTTGATAAATCAGAAAGTTATATGTGCAGATTAT TTTCCTTGGCCTTCAAGCTTCCAAAAAACTTGTAATAATCATGTTAGCTATAGCTTGTATAT ACACATAGAGATCAATTTGCCAAATATTCACAATCATGTAGTTCTAGTTTACATGCCAAAGT CTTCCCTTTTTAACATTATAAAAGCTAGGTTGTCTCTTGAATTTTGAGGCCCTAGAGATAGT CTGGCCATACCATAGATTTGGGATGATGTAGTCTGTGCTAAATATTTTGCTGAAGAAGCAGT TTCTCAGACACATCTCAGAATTTTAATTTTTAGAAATTCATGGGAAATTGGATTTTTGT AATAATCTTTTGATGTTTTAAACATTGGTTCCCTAGTCACCATAGTTACCACTTGTATTTTA AGTCATTTAAACAAGCCACGGTGGGGCTTTTTTTCTCCTCAGTTTGAGGAGAAAAATCTTGAT AATTCAAGCTGTGACTATTGTATATCTTTCCAAGAGTTGAAATGCTGGCTTCAGAATCATAC CAGATTGTCAGTGAAGCTGATGCCTAGGAACTTTTAAAGGGATCCTTTCAAAAGGATCACTT AGCAAACACATGTTGACTTTTAACTGATGTATGAATATTAATACTCTAAAAATAGAAAGACC AGTAATATAAAGTCACTTTACAGTGCTACTTCACACTTAAAAGTGCATGGTATTTTTCATG GTATTTTGCATGCAGCCAGTTAACTCTCGTAGATAGAGAAGTCAGGTGATAGATGATATTAA AAATTAGCAAACAAAAGTGACTTGCTCAGGGTCATGCAGCTGGGTGATGATAGAAGAGTGGG CTTTAACTGGCAGGCCTGTATGTTTACAGACTACCATACTGTAAATATGAGCTTTATGGTGT CATTCTCAGAAACTTATACATTTCTGCTCTCCTTTCTCCTAAGTTTCATGCAGATGAATATA AGGTAATATACTATTATATAATTCATTTGTGATATCCACAATAATATGACTGGCAAGAATTG GTGGAAATTTGTAATTAAAATAATTATTAAACCT

## FIGURE 9

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVK
LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLQPAMAV
IFSNFSIITTALLFRIVLKRRLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFF
SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFSHIRLGMGHVLIIVQCFISSMANI
YNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFYGHSAFSVALI
FVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFI
YNASKPQVPEYAPRQERIRDLSGNLWERSSGDGEELERLTKPKSDESDEDTF

### Transmembrane domains:

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293, 298-318, 328-368

#### N-glycosylation sites.

amino acids 128-132, 204-208, 218-222, 374-378

### Glycosaminoglycan attachment site.

amino acids 402-406

#### N-myristoylation sites.

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

# FIGURE 10

# FIGURE 11

CGGACGCGTGGGCGGACGCGTGGGCGGCCGGCTTGGCTAGCGCGCGGCGCCCC GTGGCTAAGGCTGCTACGAAGCGAGCTTGGGAGGAGCAGCGGCCTGCGGGGCAGAGGAGCAT CCCGTCTACCAGGTCCCAAGCGGCGTGGCCCGCGGGTCATGGCCAAAGGAGAAGGCGCCGAG AGCGGCTCCGCGGGGGGCTGCTACCCACCAGCATCCTCCAAAGCACTGAACGCCCGGCCCA GGTGAAGAAGAACCGAAAAAGAAGAACAACAGTTGTCTGTTTGCAACAAGCTTTGCTATG CTGGGATGCCATCACAGACCCCCTGGTGGGCCTCTGCATCAGCAAATCCCCCTGGACCTGCC TGGGTCGCCTTATGCCCTGGATCATCTTCTCCACGCCCCTGGCCGTCATTGCCTACTTCCTC ATCTGGTTCGTGCCCGACTTCCCACACGGCCAGACCTATTGGTACCTGCTTTTCTATTGCCT CTTTGAAACAATGGTCACGTGTTTCCATGTTCCCTACTCGGCTCTCACCATGTTCATCAGCA ACCGAGCAGACTGAGCGGGATTCTGCCACCGCCTATCGGATGACTGTGGAAGTGCTGGGCAC AGTGCTGGGCACGGCGATCCAGGGACAAATCGTGGGCCAAGCAGACACGCCTTGTTTCCAGG ACTTCAATAGCTCTACAGTAGCTTCACAAAGTGCCAACCATACACATGGCACCACTTCACAC AGGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTG TGCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGACCCTATGAAGCCCAGCAGTCTG AGCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACTT ATTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTT TTGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCT CGGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCT GTATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCCTCATGGAGAGTAA CCTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTAC TACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCAT GGAACCGAGCCCATCTTCTTCTCTCTATGTCTTCTTCACCAAGTTTGCCTCTGGAGTGTC ACTGGGCATTTCTACCCTCAGTCTGGACTTTGCAGGGTACCAGACCCGTGGCTGCTCGCAGC CGGAACGTGTCAAGTTTACACTGAACATGCTCGTGACCATGGCTCCCATAGTTCTCATCCTG CTGGGCCTGCTCTTCAAAATGTACCCCATTGATGAGGAGAGGCGGCGGCAGAATAAGAA GGCCCTGCAGGCACTGAGGGACGAGGCCAGCAGCTCTGGCTCAGAAACAGACTCCACAG AGCTGGCTAGCATCCTC**TAG**GGCCCGCCACGTTGCCCGAAGCCACCATGCAGAAGGCCACAG AAGGGATCAGGACCTGTCTGCCGGCTTGCTGAGCAGCTGGACTGCAGGTGCTAGGAAGGGAA CTGAAGACTCAAGGAGGTGGCCCAGGACACTTGCTGTGCTCACTGTGGGGCCCGGCTGCTCTG TGGCCTCCTGCCTGCCTGCCTGTGGGGCCAAGCCCTGGGGGCTGCCACTGTGAATA TTAATGTTATTAATTTTCATAAAAGCTGGAAAGC

# FIGURE 12

MWLRWALSLPPSSCLWAEPGMPSQTPWWASASANPPGPAWVALCPGSSSPRPWPSLPTSSSG SCPTSHTARPIGTCFSIASLKQWSRVSMFPTRLSPCSSATEQTERDSATAYRMTVEVLGTVL GTAIQGQIVGQADTPCFQDFNSSTVASQSANHTHGTTSHRETQKAYLLAAGVIVCIYIICAV ILILGVREQREPYEAQQSEPIAYFRGLRLVMSHGPYIKLITGFLFTSLAFMLVEGNFVLFCT YTLGFRNEFQNLLLAIMLSATLTIPIWQWFLTRFGKKTAVYVGISSAVPFLILVALMESNLI ITYAVAVAAGISVAAAFLLPWSMLPDVIDDFHLKQPHFHGTEPIFFSFYVFFTKFASGVSLG ISTLSLDFAGYQTRGCSQPERVKFTLNMLVTMAPIVLILLGLLLFKMYPIDEERRRQNKKAL QALRDEASSSGCSETDSTELASIL

# FIGURE 13

GGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTGT
GCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTGA
GCCAATCGCCTACTTCCGGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACTTA
TTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTTT
TGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCTC
GGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCTG
TATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAAC
CTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTACT
ACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCATG

# FIGURE 14

GGGGCTTCGGCCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGT ATGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAA<u>ATG</u>T GGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCT TTCATATTTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATAT CAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCCAATGCTAAATATTGCGG CAGTTTTATGCATTGCTACCATTTATGTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAA GAGAACGTTATCATCAAATTAAACAAGGCTGGCCTTGTACTTGGAATACTGAGTTGTTTAGG ACTTTCTATTGTGGCAAACTTCCAGAAAACAACCCTTTTTGCTGCACATGTAAGTGGAGCTG TGCTTACCTTTGGTATGGGCTCATTATATATGTTTGTTCAGACCATCCTTTCCTACCAAATG CAGCCCAAAATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGG AGTAAGTGCACTTAGCATGCTGACTTGCTCATCAGTTTTGCACAGTGGCAATTTTTGGGACTG ATTTAGAACAGAAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACT ACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTTTGGTTTTTTCCTGACTTACATTCGTGA TTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCTCTATGACACTG AAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGGTTCACAGAAGTTGCTTA TTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATA ATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTA TTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACTATG

# FIGURE 15

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNI
AAVLCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSG
AVLTFGMGSLYMFVQTILSYQMQPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFG
TDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLTYIRDFQKISLRVEANLHGLTLYD
TAPCPINNERTRLLSRDI

# FIGURE 16

# FIGURE 17

CCCACGCGTCCGCCGCCGCTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCCGCGGG CCGGGGTGCGAGCCGACAGGCTGGAGGCAGGTCGCTGTGGTTCCCCTCCGACCTGGCAGAGCTGCGGGAGCTCTCTGAG GTCCTTCGAGAGTACCGGAAGGAGCACCAGGCCTACGTGTTCCTGCTCTTCTGCGGCGCCCTA CCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCCTGAATGTTTTAGCTGGTGCCT TGCTACCTGCTCTCCAGTATTTTTGGCAAACAGTTGGTGGTGTCCTACTTTCCTGATAAAGT TGAGACTTTTCCCCATGACACCAAACTGGTTCTTGAACCTCTCGGCCCCAATTCTGAACATT CCCATCGTGCAGTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTTCATCTGTGT GCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTTCTCCTGGGACACTG TCTTTAAGCTGTTGGCCATTGCCATGGTGGCATTAATTCCTGGAACCCTCATTAAAAAATTT AGTCAGAAACATCTGCAATTGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAGA CACA**TGA**TCTGGATTTTCTGTTTGCCACATCCCTGGACTCAGTTGCTTATTTGTGTAATGGA TGTGGTCCTCTAAAGCCCCTCATTGTTTTTGATTGCCTTCTATAGGTGATGTGGACACTGTG CATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATCAGGT TTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTGT GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGTGATTC ACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAT ACAAAATTAGCCAGGCGTGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGC AGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGATCACACCACTGCACT CCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

# FIGURE 18

MRPLLGLLLVFAGCTFALYLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAELRELSEVLREYR KEHQAYVFLLFCGAYLYKQGFAIPGSSFLNVLAGALFGPWLGLLLCCVLTSVGATCCYLLSS IFGKQLVVSYFPDKVALLQRKVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPILNIPIVQFF FSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVFKLLAIAMVALIPGTLIKKFSQKHLQ LNETSTANHIHSRKDT

#### Important features:

### Signal peptide:

amino acids 1-17

#### Transmembrane domains:

amino acids 101-123, 189-211

#### N-glycosylation sites.

amino acids 172-176, 250-254

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 240-244, 261-265

#### N-myristoylation site.

amino acids 13-19, 104-110, 115-121, 204-210

#### Amidation site.

amino acids 27-31

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 4-15

#### Protein splicing proteins.

amino acids 25-31.

#### Sugar transport proteins.

amino acids 162-172

## FIGURE 19

CCGAGGCGGAGGCCCGAGGGGGGCCCCGCATGAATCATTGTAGTCAATCATTTT CCAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTCCGAACTGCCAGCTCAGAA TAGGAAAATAACTTGGGATTTTATATTGGAAGAC**ATG**GATCTTGCTGCCAACGAGATCAGCA TCAGAGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACC CCCCCGCAGTATCCTCTCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCT TGCTCACTGCCTACTTTGTGATTCAACCTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTCT GGAGCTCACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCCTTGCCCATTGCCAA CAGACTTTGACCCCTGGTGGACAAACGACTGTGAGCAGAATGAGTCAGAGCCCATTCCTGCC AACTGCACTGGCTGTGCCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAA ATTTGAGAGGCTCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTC AGCATTTTTTGTGCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTCGCCAAG TGGTGGCGCTGCTTTCCTGAGCGGTGGTTCCCATTTCCTTATCCATGGAGGAGACCTCTGAA CCTCTTTAAACAAGTGCTCCTTTCTTCACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAG GTGCCGAAGACATTGTCAGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTCG ACACCACCCACTGGAAGGTCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGAT  $\mathsf{GGAACCGCTTTCTCAGAACTG}$   $\mathsf{TAG}$   $\mathsf{GAAATAGAACTGTGCACAGGAACAGCTTCCAGAGCCGA$ AAACCAGGTTGAAAGGGGAAAAATAAAAACAAAAACGATGAAACTGCAAAAA

# FIGURE 20

MDLAANEISIYDKLSETVDLVRQTGHQCGMSEKAIEKFIRQLLEKNEPQRPPPQYPLLIVVY
KVLATLGLILLTAYFVIQPFSPLAPEPVLSGAHTWRSLIHHIRLMSLPIAKKYMSENKGVPL
HGGDEDRPFPDFDPWWTNDCEQNESEPIPANCTGCAQKHLKVMLLEDAPRKFERLHPLVIKT
GKPLLEEEIQHFLCQYPEATEGFSEGFFAKWWRCFPERWFPFPYPWRRPLNRSQMLRELFPV
FTHLPFPKDASLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQCRRHCQSVAMP
IEPGDIGYVDTTHWKVYVIARGVQPLVICDGTAFSEL

# FIGURE 21

## FIGURE 22

CCCACGCGTCCGCCCACGCGTCCGGCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGG CGCCAGCCATGTGGGGCCTGAGGGACACTCCCATCCCCGCTGGAAACTGTCCA GCGCCGAGACATATTCACGCATGCGTCTGAAGCTGGTGCCCAACCATCACTTCGACCCTCAC CTGGAAGCCAGCGCTCTCCGAGACAATCTGGGTGAGGTTCCCCTGACACCACCGAGGAGGC TGATGCAACTCAACACCATTGCGGGGCGGACCTACAATGACCTGTCTCAGTACCCTGTGTTC CCCTGGGTCCTGCAGGACTACGTGTCCCCAACCCTGGACCTCAGCAACCCAGCCGTCTTCCG GGACCTGTCTAAGCCCATCGGTGTGGTGAACCCCAAGCATGCCCAGCTCGTGAGGGAGAAGT GAACTTGACATGGCTGTGTCTGGATCTGAGGATGGAACTGTGATCATACACACTGTACGCCG CGGACAGTTTGTAGCGCACTACGGCCTCTGGGTGCCACATTCCCTGGACCTATTTTCCACC
TGGCATTGGGGTCCGAAGGCCAGATTGTGGTACAGAGCTCAGGGGCC
CAGGTCACCTCCTTGCACCTGTATTCAGTCAATGGGAAGTTGCGGGCTTCACTGCCCCC
GGCAGAGCAGCCTACACCCTGACGGAAGTGACAGAGGAAGTTGCTGGGCACCCCCAGGT CCTACTGAGGCGCGC**TGA**ACCTGGCCAGTCCGGCTGCTCGGGCCCCGCCCCCGGCAGGCCTGGCCCGGGAGGCCCCGCCCAGAAGTCGGCGGGAACACCCCGGGGTGGGCAGCCCAGGGGGTGA GCGGGGCCCACCCTGCCCAGCTCAGGGATTGGCGGGCGATGTTACCCCCTCAGGGATTGGCG GGCGGAAGTCCCGCCCTCGCCGGCTGAGGGGCCGCCCTGAGGGCCAGCACTGGCGTCT

# FIGURE 23

MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGL RYTAVLKQQATQHSMALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLKLVPN HHFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDELAELETP MEAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVYFYDGSTERVETEEGIGYDFRRP LAQLREVHLRRFNLRRSALELFFIDQANYFLNFPCKVGTTPVSSPSQTPRPQPGPIPPHTQV RNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGRTYNDL SQYPVFPWVLQDYVSPTLDLSNPAVFRDLSKPIGVVNPKHAQLVREKYESFEDPAGTIDKFH YGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAAWQARLESPADVKELIP EFFYFPDFLENQNGFDLGCLQLTNEKVGDVVLPPWASSPEDFIQQHRQALESEYVSAHLHEW IDLIFGYKQRGPAAEEALNVFYYCTYEGAVDLDHVTDERERKALEGIISNFGQTPCQLLKEP HPTRLSAEEAAHRLARLDTNSPSIFQHLDELKAFFAEVTVSASGLLGTHSWLPYDRNISNYF SFSKDPTMGSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWDGSLRVTALPRGKLL SQLSCHLDVVTCLALDTCGIYLISGSRDTTCMVWRLLHQGGLSVGLAPKPVQVLYGHGAAVS CVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGPIFHLALGSEGQIVVQSSA WERPGAQVTYSLHLYSVNGKLRASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA PPLPMKVAIRSVAVTKERSHVLVGLEDGKLIVVVAGQPSEVRSSQFARKLWRSSRRISQVSS GETEYNPTEAR

#### N-glycosylation site.

amino acids 677-681

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 985-989

### Tyrosine kinase phosphorylation site.

amino acids 56-65, 367-376, 543-551

#### N-myristoylation site.

amino acids 61-67, 436-442, 604-610, 610-616, 664-670, 691-697, 706-712, 711-717, 769-775, 785-791, 802-808, 820-826, 834-840, 873-879, 912-918, 954-960

## FIGURE 24

CACGGCCCACCTTGTGAACTCCTCGTGCCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCAT CCAAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGGGCTCTTC TGGACCCTTAACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTT CTACTGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCC GCACACTCCGTTACCACACTGGGTCATTGGCATTTGGAGCCCTCATCCTGACCCTTGTGCAG ATAGCCCGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGC CCGCTGCATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCC TAAACCGCAATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAA AATGCGTTCATGCTACTCATGCGAAACATTGTCAGGGTGGTCGTCCTGGACAAAGTCACAGA  $\verb|CCTGCTGCTGTTCTTTGGGAAGCTGCTGGTGGTGGAGGCGTGGGGGGTCCTGTCTTTT| \\$ TTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCCACCTCAACTATTAC TGGCTGCCCATCATGACCTCCATCCTGGGGGCCTATGTCATCGCCAGCGGCTTCTTCAGCGT TTTCGGCATGTGTGGACACGCTCTTCCTCTGCTTCCTGGAAGACCTGGAGCGGAACAACG GCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAAC GAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGGCCCTGATCCAGGACTGC ACCCCACCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGT AAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTG AGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCC GTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCAGCTAC TCGGGAGGCTGAGGCAGGAGATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGA AAAGATTTTATTAAAGATATTTTGTTAACTC

# FIGURE 25

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLF
WTLNWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQ
IARVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAK
NAFMLLMRNIVRVVVLDKVTDLLLFFGKLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYY
WLPIMTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYMSKSLLKILGKKN
EAPPDNKKRKK

## FIGURE 26

GAGTCTTGACCGCCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGT  ${\tt GGCT} {\bf ATG} {\tt TTCGTGTCCGATTTCCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCC}$ CAGTGTGACCACGTGCAATATACGCTGGTTCCAGTTTCTGGGTGGCAAGAACTTGAAACTGC ATTTCTTGAGCATAAAGAACAGTTTCATTATTTTATTCTCATAAACTGTGGAGCTAATGTAG CCAGTCAATGTCGTCAATGTATACAACGATACCCAGATCAAATTACTCATTAAACAAGATGA TGACCTTGAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATT CAGGAAATGACAGTGATGGGTCAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATA GTGGAGCAAACCATGCGGAGGAGGCAGCGGCGAGAGTGGGAGGCCCGGAGAAGAGACATCCT CTTTGACTACGAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGATGTTTGAGCTGG CTTGGATGCTGTCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGAC CAGTGGGTGCAAGACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGTCCTGCAGCG CCACGTTTCCCGCCACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCA CACGGATCTCCTTTGAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGAC AGCCTGTGCAACACCAGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAA GCGGCTCCAGGAGTTCCTTGCAGACATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAGTTCC AGGCCATGGACATCTCCTTGAAGGAGAATTTGCGGGAAATGATTGAAGAGTCTGCAAATAAA TCTGGCCAGCGACGTGTCTTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCT CAGGGACAGATCACTTCATCCAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTG TACCATGGCCTGGAACTCGCCAAGAAGCAGCTGCGAGCCACCCAGCAGACCATTGCCAGCTGC CTTTGCACCAACCTCGTCATCTCCCAGGGGCCTTTCCTGTACTGCTCTCTCATGGAGGGCAC TCCAGATGTCATGCTGTTCTCTAGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCA AGTCCTTTGTGTGTTCGACAAAGAACCGGCGCTGCAAACTGCTGCCCCTGGTGATGGCTGCC CCCCTGAGCATGGAGCATGGCACAGTGACCGTGGTGGGCATCCCCCCAGAGACCGACAGCTC GGACAGGAAGAACTTTTTTGGGAGGGCGTTTGAGAAGGCAGCGGAAAGCACCAGCTCCCGGA TGCTGCACAACCATTTTGACCTCTCAGTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTT  $\mathtt{CTGGACGCACTTATTTCCCTCCTGTCC} \underline{\mathbf{TAG}} \mathtt{GAATTTGATTCTTCCAGAATGACCTTCTTATT}$ TATGTAACTGGCTTTCATTTAGATTGTAAGTTATGGACATGATTTGAGATGTAGAAGCCATT TTTTATTAAATAAAATGCTTATTTTAGGAAA

# FIGURE 27

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETAF
LEHKEQFHYFILINCGANVDLLDILQPDEDTIFFVCDSHRPVNVVNVYNDTQIKLLIKQDDD
LEVPAYEDIFRDEEEDEEHSGNDSDGSEPSEKRTRLEEEIVEQTMRRRQRREWEARRRDILF
DYEQYEYHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGVLQRH
VSRHNHRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSLCNTSYTAARFKLWSVHGQKR
LQEFLADMGLPLKQVKQKFQAMDISLKENLREMIEESANKFGMKDMRVQTFSIHFGFKHKFL
ASDVVFATMSLMESPEKDGSGTDHFIQALDSLSRSNLDKLYHGLELAKKQLRATQQTIASCL
CTNLVISQGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKSFVCSTKNRRCKLLPLVMAAP
LSMEHGTVTVVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNHFDLSVIELKAEDRSKFL
DALISLLS

# FIGURE 28

# FIGURE 29

CACGAAGGGACGATACCAGAAAACACCCTCAACCCAAAGGAAATAGACTACAGCCCCAATTG GTCTTCCCTTTATCGAGTCAAGAAACCCCCCCTTCTTGAGCTATTTACAGCTTTTAACAATT GGCGAGGACTTCCAGACGCTGCTCTTCGAAAGGGTCTTTGTGAACCTTGATGGGTGCTTTGA CATGGCGACCGGCCAGTTTGCTGCTCCCCTGCGTGGCATCTACTTCTTCAGCCTCAATGTGC ACAGCTGGAATTACAAGGAGACGTACGTGCACATTATGCATAACCAGAAAGAGGCTGTCATC CTGTACGCGCAGCCAGCGAGCGCAGCATCATGCAGAGCCAGAGTGTGATGCTGGACCTGGCCTACGGGGACCGCGTCTGGGTGCGGCTCTTCAAGCGCCAGCGCGAGAACGCCATCTACAGCA AGCACTTCTCAAACTTGGAAATGCATGCGAATCACCCGGGGTTCGTGTTAAATGCAGATTCT GACTCAGCAGGTCTGAGTGGGTCCAGGATTCTGTGTTTCTCATATGTTCCTGGGTGATGCTG ACCCAGGAGTGTGGGGGCATTTGGGGGGGTGAAGTGGCCCCCGAAGAATGGAACCCACACCCA TAGCTCTCCCCACAGCTGATACGGCATCCTGCGAGAAGACCTGCCCTCCTCACTGGGATCCC CTTCCTGCCTCCCAGGGCTCTGCCAGGGCCTTGCTCAGTCCCTTCCACCAAAGTCATCT 

# FIGURE 30

MVTAALGPVWAALLIFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSGRPH ALPEIRPYINITILKGDKGDPGPMGLPGYMGREGPQGEPGPQGSKGDKGEMGSPGAPCQKRF FAFSVGRKTALHSGEDFQTLLFERVFVNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKET YVHIMHNQKEAVILYAQPSERSIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYIT FSGHLIKAEDD

### Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-75

Clq domain proteins.

amino acids 144-178, 78-111 and 84-117

## FIGURE 31

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCTCGGGCCCGACCCGCCAGGAAAGACTG CTCCAGGGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGG GCTGCCCATCCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGG ACCACGGTGCCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCAT CACCATGCTCGACGCAGGCAGCTTTGCCGGCCTGCCGGGCCTGCAGCTCCTGGACCTGTCAC AGAACCAGATCGCCAGCCTGCCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTG GACCTGACGGCCAACAGGCTGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCCT CGAGCGCCTCTACCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGC TCGACCGCCTCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGC CTGCCCGCCTGCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCAT CCTGGACACTGCCAACGTGGAGGCGCTGCGGCTGGTCTGGGGCTGCAGCAGCTGGACG AGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACCTGGATGTGTCCGACAACCAGCTGGAG CGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACGCCCTGCGGCTGGCCGGCAACAC CCGCATTGCCCAGCTGCGGCCGAGGACCTGGCCGGCCTGGCTGCCCTGCAGGAGCTGGATG TGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCCTCTTCCCCCGCCTGCGG CTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGCTGGTTTGGCCCCTG GGTGCGCGAGAGCCACGTCACACTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCCGCCCA AGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCCAGCCACCACC ACCACAGCCACAGTGCCCACCACGAGGCCCGTGGTGCGGGAGCCCACAGCCTTGTCTTCTAG CTGCCCCACCGACTGTAGGGCCTGTCCCCCAGCCCCAGGACTGCCCACCGTCCACCTGCCTC AATGGGGGCACATGCCACCTGGGGACACGGCACCACCTGGCGTGCTTGTGCCCCGAAGGCTT CACGGGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCCAGCCCTACACCAGTCA CGCCGAGGCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGC GTGGGGCTGCAGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTA AGTACACGGTCACCCAGCTGCGGCCCAACGCCACTTACTCCGTCTGTGTCATGCCTTTGGGG CCCGGGCGGTGCCGAGGGCGAGGAGGCCTGCGGGGAGGCCCATACACCCCCAGCCGTCCA CTCCAACCACGCCCAGTCACCCAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCG CCCTGGCCGCGGTGCTCCTGGCCGCGCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGG GGGCGGCCATGGCAGCAGCGCTCAGGACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCCT GGAACTGGAGGGAGTGAAGGTCCCCTTGGAGCCAGGCCCGAAGGCAACAGAGGCGGTGGAG AGGCCCTGCCCAGCGGGTCTGAGTGTGAGGTGCCACTCATGGGCTTCCCAGGGCCTGGCCTC CAGTCACCCCTCCACGCAAAGCCCTACATC**TAA**GCCAGAGAGAGACAGGGCAGCTGGGGCCG GGCTCTCAGCCAGTGAGATGGCCAGCCCCTCCTGCTGCCACACCACGTAAGTTCTCAGTCC CAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCTGGGCCCTGTTCCCTCTGGA CCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCCAGCTGACGAGCCCTAACGTCCCCAGAAC CGAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTCCCTGGGCACGGCG GGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCCTGCTGGGCTCTCCCACTCCAGGCGGA CCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGCGGCTGTG TGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAAACTGGAAAGGAAGATGCTTTA GGAACATGTTTTGCTTTTTTAAAATATATATATTTTATAAGAGATCCTTTCCCATTTATTCTG GGAAGATGTTTTTCAAACTCAGAGACAAGGACTTTGGTTTTTGTAAGACAAACGATGATATG AAGGCCTTTTGTAAGAAAAAATAAAAGATGAAGTGTGAAA

## FIGURE 32

MCSRVPLLLPLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFEN
GITMLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLR
RLERLYLGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEP
GILDTANVEALRLAGLGLQQLDEGLFSRLRNLHDLDVSDNQLERVPPVIRGLRGLTRLRLAG
NTRIAQLRPEDLAGLAALQELDVSNLSLQALPGDLSGLFPRLRLLAAARNPFNCVCPLSWFG
PWVRESHVTLASPEETRCHFPPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALS
SSLAPTWLSPTAPATEAPSPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPE
GFTGLYCESQMGQGTRPSPTPVTPRPPRSLTLGIEPVSPTSLRVGLQRYLQGSSVQLRSLRL
TYRNLSGPDKRLVTLRLPASLAEYTVTQLRPNATYSVCVMPLGPGRVPEGEEACGEAHTPPA
VHSNHAPVTQAREGNLPLLIAPALAAVLLAALAAVGAAYCVRRGRAMAAAAQDKGQVGPGAG
PLELEGVKVPLEPGPKATEGGGEALPSGSECEVPLMGFPGPGLQSPLHAKPYI

## FIGURE 33

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTGAGCTCAT CTTCATCATTCATATGAGGAAATAAGTGGTAAAATCCTTGGAAATACA**ATG**AGACTCATCAG  $\mathtt{AAACATTTACATATTTTGTAGTATTGTTATGACAGCAGAGGGTGATGCTCCAGAGCTGCCAG}$ ATCTCAAAACCTTTGAATTCAACAAGGAGTTAAGATATTTAGATTTGTCTAATAACAGACTG GTGGGGCAAAAATACAAAAATCAGATTTCCAGAAAATTGCTCATCTGCATCTAAATACTGTC TTCTTAGGATTCAGAACTCTTCCTCATTATGAAGAAGGTAGCCTGCCCATCTTAAACACAAC AAAACTGCACATTGTTTTACCAATGGACACAAATTTCTGGGTTCTTTTGCGTGATGGAATCA AGACTTCAAAAATATTAGAAATGACAAATATAGATGGCAAAAGCCAATTTGTAAGTTATGAA ATGCAACGAAATCTTAGTTTAGAAAATGCTAAGACATCGGTTCTATTGCTTAATAAAGTTGA ACAGGATAAAATCTATTTGCTTTTGACCAAAATGGACATAGAAAACCTGACAATATCAAATG CACAAATGCCACACATGCTTTTCCCGAATTATCCTACGAAATTCCAATATTTAAATTTTGCC AATAATATCTTAACAGACGAGTTGTTTAAAAGAACTATCCAACTGCCTCACTTGAAAACTCT CATTTGAATGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACACCCCTTGGAACACTTGAGTCAAAATCTATTACAACATAAAAATGATGAAAAATTGCTCA TGGCCAGAAACTGTGGTCAATATGAATCTGTCATACAATAAATTGTCTGATTCTGTCTTCAGGTGCTTGCCCAAAAGTATTCAAATACCTAAATAATAACCAAAATCCAAACTGTACCTA AAGAGACTATTCATCTGATGGCCTTACGAGAACTAAATATTGCATTTAATTTTCTAACTGAT ATGATGGTTGGATGCTGTGAATTAAAAAATTTCATTCAGCTTGAAACATATTCAGAGGTC
ATGATGGTTGGATGGTCAGATTCATACACCTGTGAATACCCTTTAAACCTAAGGGGAACTAG
GTTAAAAGACGTTCATCTCCACGAATTATCTTGCAACACACGCTCTGTTGATTGTCACCATTG
TGGTTATTATGCTAGGTTCTGGGGTTGGCTTGTGCCCTTCTCCACTTTGATCTCCCC
TGGTATCTCAGGATGCTAGGTCAATGCACACAAACATGGCACAGGGTTAGGAAAACAACCCA
AGAACAACTCAAAGAAATGTCCGATTCCACGCATTTCATACAGTGAACATGATCTC ACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCAGAGAAATGTATGAACTGCAGACA TTCACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCTCTCTGATGAGAACAGATTGTCT A**TAA**AATCCCACAGTCCTTGGGAAGTTGGGGACCACATACACTGTTGGGATGTACATTGATA CAACCTTTATGATGGCAATTTGACAATATTTATTAAAATAAAAAATGGTTATTCCCTTCATA TCAGTTTCTAGAAGGATTTCTAAGAATGTATCCTATAGAAACACCTTCACAAGTTTATAAGG GCTTATGGAAAAAGGTGTTCATCCCAGGATTGTTTATAATCATGAAAAATGTGGCCAGGTGC TCATGGCCACAAAATAAGGTCTAATTCAATAAATTATAGTACATTAATGTAATATAATATTA CATGCCACTAAAAAAAATAATGTTATATTTCCTGGTATGGAAAAAACATATTAATAT CATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCCTGGTATGGAAAAAACATATTAATAT GTTATAAACTATTAGGTTGGTGCAAAACTAATTGTGGTTTTTGCCATTGAAATGGCATTGAA ATAAAAGTGTAAAGAAATCTATACCAGATGTAGTAACAGTGGTTTTGGGTCTGGGAGGTTGGA TTACAGGGAGCATTTGATTTCTATGTTGTATTTCTATAATGTTTGAATTGTTTAAATGA ATCTGTATTTCTTTTATAAGTAGAAAAAAAATAAAGATAGTTTTTACAGCCT

# FIGURE 34

MRLIRNIYIFCSIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQ
LQSSDFHSVSKLRVLILCHNRIQQLDLKTFEFNKELRYLDLSNNRLKSVTWYLLAGLRYLDL
SFNDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLP
ILNTTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLL
LNKVDLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFR
VFYIQQDKIYLLLTKMDIENLTISNAQMPHMLFPNYPTKFQYLNFANNILTDELFKRTIQLP
HLKTLILNGNKLETLSLVSCFANNTPLEHLDLSQNLLQHKNDENCSWPETVVNMNLSYNKLS
DSVFRCLPKSIQILDLNNNQIQTVPKETIHLMALRELNIAFNFLTDLPGCSHFSRLSVLNIE
MNFILSPSLDFVQSCQEVKTLNAGRNPFRCTCELKNFIQLETYSEVMMVGWSDSYTCEYPLN
LRGTRLKDVHLHELSCNTALLIVTIVVIMLVLGLAVAFCCLHFDLPWYLRMLGQCTQTWHRV
RKTTQEQLKRNVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENI
VSFIEKSYKSIFVLSPNFVQNEWCHYEFYFAHHNLFHENSDHIILILLEPIPFYCIPTRYHK
LKALLEKKAYLEWPKDRRKCGLFWANLRAAINVNVLATREMYELQTFTELNEESRGSTISLM

## FIGURE 35

GGGGGCTTTCTTGGGCTTGGCTTGGAACACCTGCCTCCAAGGACCGGCCTCGGAGGGGTCGCCGGGAAAGG GAGGGAAGAAGGAAGGCCGGCCCGCCCCTGCGCCCCCGCGCCCTCTGCGCCCCTGTCCGCCCCGGC  $\tt CTGGCAGTGACCCTGGCCGGGGTCGGAGCCCAGGGCGCAGCCCTCGAGGACCCTGATTATTACGGGCAGGAGAT$ AAGAGGGAGAAGTCGGCTCCGGAGCCGCCTCCACCAGGTAAACACAGCAACAAAAAAGTTATGAGAACCAAGAG TTGGTCTGGAAACCTTAAAAATCACAGACTTCCAGCTCCATGCCTCCACGGTGAAGCGCTATGGCCTGGGGGCA CATCGAGGGAGACTCAACATCCAGGCGGGCATTAATGAAAATGATTTTTATGACGGAGCGTGGTGCGCGGGAAG AAATGACCTCCAGCAGTGGATTGAAGTGGATGCTCGGCGCCTGACCAGATTCACTGGTGTCATCACTCAAGGGA GGAACTCCCTCTGGCTGAGTGACTGGGTGACATCCTATAAGGTCATGGTGAGCAATGACAGCCACACGTGGGTC ACTGTTAAGAATGGATCTGGAGACATGATATTTGAGGGAAACAGTGAGAAGGAGATCCCTGTTCTCAATGAGCT  ${\tt ACCCGTCCCCATGGTGGCCCGCTACATCCGCATAAACCCTCAGTCCTGGTTTGATAATGGGAGCATCTGCATGA}$ GAATGGAGATCCTGGGCTGCCCACTGCCAGATCCTAATAATTATTATCACCGCCGGAACGAGATGACCACCACT GATGACCTGGATTTTAAGCACCACAATTATAAGGAAATGCGCCAGTTGATGAAAGTTGTGAAATGTGTCC CAATATCACCAGAATTTACAACATTGGAAAAAGCCACCAGGGCCTGAAGCTGTATGCTGTGGAGATCTCAGATC ACCCTGGGGAGCATGAAGTCGGTGAGCCCGAGTTCCACTACATCGCGGGGGCCCCACGGCAATGAGGTGCTGGGC  $\tt CGGGAGCTGCTGCTGCTGGTGCAGTTCGTGTCAGGAGTACTTGGCCCGGAATGCGCGCATCGTCCACCT$ GGTGGAGGAGACGCGGATTCACGTCCTCCCTCCCTCAACCCCGATGGCTACGAGAAGGCCTACGAAGGGGGGCT CGGAGCTGGGAGGCTGGTCCCTGGGACGCTGGACCCACGATGGAATTGACATCAACAACAACTTTCCTGATTTA AACACGCTGCTCTGGGAGGCAGAGGATCGACAGAATGTCCCCAGGAAAGTTCCCAATCACTATATTGCAATCCC ACACCGCCTCATGACAGACGCCCGGAGGAGGGTGTGCCACACGGAGGACTTCCAGAAGGAGGAGGCACTGTCA ATGGGGCCTCCTGGCACCCGTCGCTGGAAGTCTGAACGATTTCAGCTACCTTCATACAAACTGCTTCGAACTG TCCATCTACGTGGGCTGTGATAAATACCCACATGAGAGCCAGCTGCCCGAGGAGTGGGAGAATAACCGGGAATC  ${\tt TCTGATCGTGTCATGGAGCAGGTTCATCGTGGCATTAAAGGCTTGGTGAGAGATTCACATGGAAAAGGAATCC}$ CAAACGCCATTATCTCCGTAGAAGGCATTAACCATGACATCCGAACAGCCAACGATGGGGATTACTGGCGCCTC CTGAACCCTGGAGAGTATGTGGTCACAGCAAAGGCCGAAGGTTTCACTGCATCCACCAAGAACTGTATGGTTGG CTATGACATGGGGGCCACAAGGTGTGACTTCACACTTAGCAAAACCAACATGGCCAGGATCCGAGAGATCATGG TGGACTCACTCACTGTTGTTTCCTCTGTAATTCAAGAAGTGCCTGGAAGAGGGGTGCATTGTGAGGCAGGTCC CAAAAGGGAAGGCTGGAGGCTGTTTTCTTTTCTTTGTTCCCATTTATCCAAATAACTTGGACAGAGCA GAGCCTGTCCGTTCAGAGCCTCTGGCTGCATAGAAAAGGATTCTGGTGCTTCCCCTGTTTGCGTGGCAGCAAGG GTTCCACGTGCATTTGCAATTTGCACAGCTAAAATTGCAGCATTTCCCCAGCTGGGCTGTCCCAAATGTTACCA TTTGAGATGCTCCCAGGCGTCCTAAGAGAATCCACCCTCTCTGGCCCTGGGACATTGCAAGCTGCTACAAATAA ATTCTGTGTTCTTTTGACAATAGCGTCATTGCCAAGTGCACATCAGTGAGCCTCTTGAATCTGTTTAGTCTCCT TGGAGCTTCTTGCACAAATTCTGGGTCCATAAACAACCCCCAAAGTCCCTGCTGATCCAGTAGCCCTGGAGGTT GAATTGAGTGCTCATGGGTTGGCCTCATATCAGCCTGGGAGTTATTTTTGATATGTAGAATGCCAGATCTTCCA GATTAGGCTAAATGTAATGAAAACCTCTTAGGATTATCTGTGGAGCATCAGTTTGGGAAGAATTATTGAATTAT 

## FIGURE 36

MSRPGTATPALALVLLAVTLAGVGAQGAALEDPDYYGQEIWSREPYYARPEPELETFSPPLP
AGPGEEWERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHS
VRVAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCA
GRNDLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSHTWVTVKNGSGDMIF
EGNSEKEIPVLNELPVPMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTT
TDDLDFKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEF
HYIAGAHGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNPDGYEKAYEGG
SELGGWSLGRWTHDGIDINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAA
ETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSYAST
HRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHES
QLPEEWENNRESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRLL
NPGEYVVTAKAEGFTASTKNCMVGYDMGATRCDFTLSKTNMARIREIMEKFGKQPVSLPARR
LKLRGRKRRQRG

### FIGURE 37

ATTTGGGGGATGTGGGACCTCCAATTCCCAGCCCCGGCTTCAGCTCTTTCCCAGGTGTTGACTCCAGCTCCAGC TTCAGCTCCAGCTCCAGGTCGGGCTCCAGCTCCAGCCGCAGCTTAGGCAGCGGAGGTTCTGTGTCCCAGTTGTT TTCCAATTTCACCGGCTCCGTGGATGACCGTGGGACCTGCCAGTGCTCTGTTTCCCTGCCAGACACCACCTTTC GTGAGGGAATATGTCCAATTAATTAGTGTGTATGAAAAGAAACTGTTAAACCTAACTGTCCGAATTGACATCAT GGAGAAGGATACCATTTCTTACACTGAACTGGACTTCGAGCTGATCAAGGTAGAAGTGAAGGAGATGGAAAAAC TGGTCATACAGCTGAAGGAGAGTTTTGGTGGAAGCTCAGAAATTGTTGACCAGCTGGAGGTGGAGATAAGAAAT ATGACTCTCTTGGTAGAGAAGCTTGAGACACTAGACAAAAACAATGTCCTTGCCATTCGCCGAGAAATCGTGGC TCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAAACACCCCTGTCGTCCACCCTCCCCACTC CAGGGAGCTGTGGTCATGGTGGTGTGAACATCAGCAAACCGTCTGTGGTTCAGCTCAACTGGAGAGGGTTT ATTGAATACAGATGGGAGACTGTTGGAGTATTATAGACTGTACAACACTGGATGATTTGCTATTGTATATAA ATGCTCGAGAGTTGCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAAC ATGTACAACACCGGGAATATTGCCAGAGTTAACCTGACCACCAACACGATTGCTGTGACTCAAACTCTCCCTAA TGCTGCCTATAATAACCGCTTTTCATATGCTAATGTTGCTTGGCAAGATATTGACTTTGCTGTGGATGAGAATG GATTGTGGGTTATTTATTCAACTGAAGCCAGCACTGGTAACATGGTGATTAGTAAACTCAATGACACCACACTT GTATGCCACCCGTACTATGAACACCAGAACAGAAGAGATTTTTTACTATTATGACACAAACACAGGGAAAGAGG GCAAACTAGACATTGTAATGCATAAGATGCAGGAAAAAGTGCAGAGCATTAACTATAACCCTTTTGACCAGAAA  $\texttt{CTTTATGTCTATAACGATGGTTACCTTCTGAATTATGATCTTTCTGTCTTGCAGAAGCCCCAG} \underline{\textbf{TAA}} \texttt{GCTGTTTA}$  $\tt CTAAAAGTGTGTTCATTTTGCAGCAATGTTTAGGTGCATAGTTCTACCACACTAGAGATCTAGGACATTTGTCT$ TTGTCAGAGGTCTAGGGGCACTGTGGGCCTAGTGAAGCCTACTGTGAGGAGGCTTCACTAGAAGCCTTAAATTA GGAATTAAGGAACTTAAAACTCAGTATGGCGTCTAGGGATTCTTTGTACAGGAAATATTGCCCAATGACTAGTC GGAGCTCCTCGAGGGACCAAATCTCCAACTTTTTTTCCCCTCACTAGCACCTGGAATGATGCTTTGTATGTGG CAGATAAGTAAATTTGGCATGCTTATATATTCTACATCTGTAAAGTGCTGAGTTTTATGGAGAGAGGCCTTTTT ATGCATTAAATTGTACATGGCAAATAAATCCCAGAAGGATCTGTAGATGAGGCACCTGCTTTTTCTTTTCTCTC AACCAGACTTACTAACCAATTCCACCCCCACCAACCCCCTTCTACTGCCTACTTTAAAAAAATTAATAGTTTT AGACTATAAGAAAATCTGATGGCAGTGACAAAGTGCTAGCATTTATTGTTATCTAATAAAGACCTTGGAGCATA TGTGCAACTTATGAGTGTATCAGTTGTTGCATGTAATTTTTGCCTTTGTTTAAGCCTGGAACTTGTAAGAAAAT GAAAATTTAATTTTTTTTTTTCTAGGACGAGCTATAGAAAAGCTATTGAGAGTATCTAGTTAATCAGTGCAGTAGT TGGAAACCTTGCTGGTGTATGTGATGTGCTTCTGTGCTTTTGAATGACTTTATCATCTAGTCTTTGTCTATTTT 

## FIGURE 38

# FIGURE 39

GCTCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAACACCCCTGTCGTCCAC
CCTCCTCCCACTCCAGGGAGCTGTGGTCATGGTGGTGGTGAACATCAGCAAACCGTCTGT
GGTTCAGCTCAACTGGAGAGGGTTTTCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTC
CCCAGCATCCAAACAAAGGNATGTATTGGGNGGCGCCATTGAATACAGATGGGAGACTGTTG
GAGTATTATAGACTGTACAACCCACTGGATGATTTGCTATTGTATATAAATGCTCGAGAGTT
GCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAACA

### FIGURE 40

 $\texttt{CCGCTGCTCTTGTGACGTTGTGGAG} \textbf{\underline{ATG}} \textbf{GGGAGCGTCCTGGGGGCTGTGCTCCATGGCGAGCT}$ TGTAATGTTGATACCAGGAATGGAAGAACAACTGAATAAGATTCCTGGATTTTGTGAGAATG AGAAAGGTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTT TATCAGCTACAGCTCTGAATTATCTGCTGTCTTTAGTTGCTATCGTCCTGTTCTTTGTCTAC TACACTCATCCAGCCAGTTGTTCAGAAAACAAGGCGTTCATCAGTGTCAACATGCTCCTCTG CGTTGGTGCTTCTGTAATGTCTATACTGCCAAAAATCCAAGAATCACAACCAAGATCTGGTT TGTTACAGTCTTCAGTAATTACAGTCTACACAATGTATTTGACATGGTCAGCTATGACCAAT GAACCAGAAACAAATTGCAACCCAAGTCTACTAAGCATAATTGGCTACAATACAACAAGCAC TGTCCCAAAGGAAGGGCAGTCAGTCCAGTGGTGGCATGCTCAAGGAATTATAGGACTAATTC TCTTTTTGTTGTGTGTATTTTATTCCAGCATCCGTACTTCAAACAATAGTCAGGTTAATAAA CTGACTCTAACAAGTGATGAATCTACATTAATAGAAGATGGTGGAGCTAGAAGTGATGGATC ACTGGAGGATGGGGACGATGTTCACCGAGCTGTAGATAATGAAAGGGATGGTGTCACTTACA GTTATTCCTTCTTTCACTTCATGCTTTTCCTGGCTTCACTTTATATCATGATGACCCTTACC AACTGGTCCAGGTATGAACCCTCTCGTGAGATGAAAAGTCAGTGGACAGCTGTCTGGGTGAAAATCTCTTCCAGTTGGATTGGCATCGTGCTGTATGTTTGGACACTCGTGGCACCACTTGTTC GCAAATTTGTTTGACTATTTTAAATATTTTGATGTTTGCCTGGCAGGATACTGCAAAGAAC ATGGTTTATTTTAAAATTTATAAACAAGTCACTTAAATGCCAGTTGTCTGAAAAATCTTATA AGGTTTTACCCTTGATACGGAATTTACACAGGTAGGGAGTGTTTAGTGGACAATAGTGTAGG TTATGGATGGAGGTGTCGGTACTAAATTGAATAACGAGTAAATAATCTTACTTGGGTAGAGA TGGCCTTTGCCAACAAAGTGAACTGTTTTGGTTGTTTTAAACTCATGAAGTATGGGTTCAGT GGAAATGTTTGGAACTCTGAAGGATTTAGACAAGGTTTTGAAAAGGATAATCATGGGTTAGA AGGAAGTGTTTTGAAAGTCACCTTTGAAAGTTAGTTTTTGGGCCCAGCACGGTAGCTCACCCTT GGTAATCCCAGCACTTTGGGAGCTTAAGTGGGTAGATTACTTGAGCCCAGGAATTCAGACCA GCTTGGCACATGGTGAACCTGTTCTATAAAAATAATCTGGCTTTGAGCATATGCCTGTGGTC CAGCACTGAGAGGCTAGTGAAGATTGCTGAGCCCAGAGCCAAAGGTTGCAGTGAGCAAGTCA AGGCAAAATTTTGACAGGGAAGGAAGTAACTGCAAAACCACTAGGCTTTAGTAGGTACTTAT ATAAAATCTAGTCCAGTTCTCATTTAAAAAAATGAAGACACTGAAATACAGACTTAAATA GATGTGGATTGCTGGTGTCCAGCATGACCCATAAACAGGTCAGAAGAATGATGGAATGTTTT AGAATAAACTCCTGCTTATAGTATACTACACAGTTCAAAAGATGTTTAAAATGCTTTTGTAT TTACTGCCATGTAATTGAAATATATAGATTATTGTAACCTTTCAACCTGAAAATCAAGCAGT ATGAGAGTTTAGTTATTTGTATGTGTCACTAGTGTCTAATGAAGCTTTTAAAATCTACAATT TCTTCTTTAAAATATTTTATTTGTGAATGGAATATAACAATTCAGCTTAATTCCCCAACC TTATTCTGTGTGTAGACATTGTATTCCACAATTTTGAATGGCTGTGTTTTACCTCTAAATAA ATGAATTCAGAGAAAAAAAAAAAAAAAA

# FIGURE 41

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGME
EQLNKIPGFCENEKGVVPCNILVGYKAVYRLCFGLAMFYLLLSLLMIKVKSSSDPRAAVHNG
FWFFKFAAAIAIIIGAFFIPEGTFTTVWFYVGMAGAFCFILIQLVLLIDFAHSWNESWVEKM
EEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPASCSENKAFISVNMLLCVGASVMSI
LPKIQESQPRSGLLQSSVITVYTMYLTWSAMTNEPETNCNPSLLSIIGYNTTSTVPKEGQSV
QWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTSDESTLIEDGGARSDGSLEDGDDVH
RAVDNERDGVTYSYSFFHFMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVWVKISSSWIGI
VLYVWTLVAPLVLTNRDFD

## FIGURE 42

GCGAGAAAGAAGCTGTCTCCATCTTGTCTGTATCCCGCTGCTTCTTGNGACGTTGTGGAGAT
GGGGAGCGTCCCTGGGGCTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCC
CCGTGTTTGCTATGCCGATGCTGTCCTAGTGGAAACAANTCCACTGTAACTAGATTGATCTA
TGCACTTTTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAG
AACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATT
TTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTTGGCTATGTTCTATCTTCT
CTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGAT
TTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

## FIGURE 43

GTTATTGTGAACTTTGTGGAGATGGGAGGTCNTGGGGCTGTTTCCATGGCGAGCTGGATAC
CANGTTTGTGTGGAAGTGCCCCGTGTTTGNTATGCCGATGCTGTCCTAGTGGAAACAANTCC
ACTGTAATTAGATTGATNTATGCACTTTTNTTGCTTGTTGGAGTANGTGTAGCTTGTGTAAT
GTTGATACCAGGAATGGAAGAACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAG
GTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATNGTTTGTGCTTTGGTTTG
GCTANGTTCTATNTTCTTCTCTCTTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAG
AGCTGCAGTGCACAATGGATTTTGGTTTTTAAATTTGCTGCAGCAATTGCAATTATTATTG

## FIGURE 44

Hinto Albahil dah

## FIGURE 45

## FIGURE 46

GGCAAATGTGTGTGGCTGGAGGCGAGCGCGAGGCTTTCGGCAAAGGCAGTCGAGTGTTTGCAGACCGGGGCGAG GTCGTTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTAGCGATGCC CCCTGGTTTGTGTGTTACGCACACACGTGCACACAAGGCTCTGGCTCGCTTCCCTCCTCGTTTCCAGCTCC TGGGCGAATCCCACATCTGTTTCAACTCTCCGCCGAGGGCGAGCAGGAGCGAGAGTGTGTCGAATCTGCGAGTG AAGAGGGACGAGGGAAAAGAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAAAGAAGCACCAGAT TGGAAGCTCGGCCTTCCTGTCGCACCACCGCCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCCGCCCCA ACATCATCCTGGTGCTGACGGACGACCAGGATGTGGAGCTGGGTTCCATGCAGGTGATGAACAAGACCCGGCGC ATCATGGAGCAGGGGGGGGCGCACTTCATCAACGCCTTCGTGACCACACCCATGTGCTGCCCCTCACGCTCCTC CATCCTCACTGGCAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGC AGGCACAGCACGAGAGCCGCACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAG TATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTCGGACTCCTTAAAAACTCCCG CTTTTATAACTACACGCTGTGTCGGAACGGGGTGAAAGAGAAGCACGGCTCCGACTACTCCAAGGATTACCTCA CAGACCTCATCACCAATGACAGCGTGAGCTTCTTCCGCACGTCCAAGAAGATGTACCCGCACAGGCCAGTCCTC ATGGTCATCAGCCATGCAGCCCCCACGGCCCTGAGGATTCAGCCCCACAATATTCACGCCTCTTCCCAAACGC ATCTCAGCACATCACGCCGAGCTACAACTACGCGCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGC  ${\tt CCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGCAGACCCTCATGTCGGTGGACCCTCATGTCATGTCGGTGGACCCTCATGTCATGTCGGTGGACCCTCATGTCATGTCGGTGGACCCTCATGTCATGTCGGTGGACCCTCATGT$ GACTCCATGGAGACGATTTACAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGC CGACCACGGTTACCACATCGGCCAGTTTGGCCTGGTGAAAGGGAAATCCATGCCATATGAGTTTGACATCAGGG TCCCGTTCTACGTGAGGGGCCCCAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCCTCAACATTGACCTG GGACACGGAGCGGCCGGTGAATCGGTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGG AGAGAGGCAAGCTGCTACACAAGAGAGACAATGACAAGGTGGACGCCCAGGAGGAGAACTTTCTGCCCAAGTAC CAGCGTGTGAAGGACCTGTGTCAGCGTGCTGAGTACCAGACGGCGTGTGAGCAGCTGGGACAGAAGTGGCAGTG TGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCC TCTCCAACCTCGTGCCCAAGTACTACGGGCAGGGCAGCGAGGCCTGCACCTGTGACAGCGGGGGACTACAAGCTC AGCCTGGCCGGACGCCGGAAAAAACTCTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCAGTCGCTCCATCCG CTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCCAGCCCCGAAACCTCA CCAAGCGGCACTGGCCAGGGGCCCCTGAGGACCAAGATGACAAGGATGGTGGGGACTTCAGTGGCACTGGAGGC CTTCCCGACTACTCAGCCGCCAACCCCATTAAAGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCCA GTGTGACCTGGACCTGTACAAGTCCCTGCAGGCCTGGAAAGACCACAAGCTGCACATCGACCACGAGATTGAAA CCCTGCAGAACAAATTAAGAACCTGAGGGAAGTCCGAGGTCACCTGAAGAAAAAGCGGCCAGAAGAATGTGAC TGTCACAAAATCAGCTACCACACCCAGCACAAAGGCCGCCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTCAG GAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAGAAAACTCCGCAAGCTGC TCAAGCGCCTGCAGAACAACGACACGTGCAGCATGCCAGGCCTCACGTGCTTCACCCACGACAACCAGCACTGG CAGACGGCGCCTTTCTGGACACTGGGGCCTTTCTGTGCCTGCACCAGCGCCAACAATAACACGTACTGGTGCAT GAGGACCATCAATGAGACTCACAATTTCCTCTTCTGTGAATTTGCAACTGGCTTCCTAGAGTACTTTGATCTCA ACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACTGGACAGGGATGTCCTCAACCAGCTACACGTACAG AAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGG  ${\tt GACAACTGTGGGGAAGGTTAA}{\tt GAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACCTGA}$  $\tt CTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCCTGAGAAAGC$ AAGCACGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCAGCAGGAGCAGAGATAACTTCAGGAAGTCC ATTTTTGCCCCTGCTTTTGCTTTGGATTATACCTCACCAGCTGCACAAAATGCATTTTTTCGTATCAAAAAGTC TCCCAAGGGCGAAAGTCATTGGAATTTTTAAATCATAGGGGAAAAGCAGTCCTGTTCTAAATCCTCTTATTCTT TTGGTTTGTCACAAAGAAGGAACTAAGAAGCAGGACAGAGGCAACGTGGAGAGGCTGAAAACAGTGCAGAGACG TTTGACAATGAGTCAGTAGCACAAAAGAGATGACATTTACCTAGCACTATAAACCCTGGTTGCCTCTGAAGAAA  $\verb|CTGCCTTCATTGTATATATGTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGGAACCTAATAAGAAATCAACATGGGAACTTTTAGGGGAACCTAATAAGAAATCAACATGGGAACTTTTAGGGGAACCTAATAAGAAATCAACATGGGAACTTTTAGGGGAACCTAATAAGAAATCAACATGGGAACTTTTAGGGGAACCTAATAAGAAATCAACATGGGAACTTTTAGGGGAACCTAATAAGAAATCAACATGGGAACTTTTAGGGGAACCTAATAAGAAATCAACATGGGAACTTTTAGGGGAACCTAATAAGAAATCAACATGGGAACTTTTAGGGGAACCTAATAAGAAATCAACATGGGAACTTTTAGGGGGAACCTAATAAGAAATCAACATGGGAACTTTTAGGGGGAACCTAATAAGAAATTAGAACATGGGAACTTTTAGGGGGAACCTAATAAGAAATTAGAACATGGGAACTTTTAGGGGGAACCTAATAAGAAATTAGAACATGGGAACTTTTAGGGGGAACCTAATAAGAAATTAGAACATGGGAACTTTTAGGGGGAACCTAATAAGAAATTAGAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAAATGAAATGAAAATGAAAATGAAAATGAAATGAAAATGAAAATGAAAATGAATGAAATGAAATTAAATGAAATGAAATGAAATGAATTTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG$ CCCAATTTTCAGGAGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAAGAAAAA

## FIGURE 47

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNIILVLTDDQDVELGSMQ
VMNKTRRIMEQGGAHFINAFVTTPMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQAQHES
RTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSD
YSKDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFPNASQHITP
SYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNT
YIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDI
AGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEEN
FLPKYQRVKDLCQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKY
YGQGSEACTCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSVAIEVDGRVYHVGLGDAAQ
PRNLTKRHWPGAPEDQDDKDGGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKS
LQAWKDHKLHIDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGRLKHRGSSL
HPFRKGLQEKDKVWLLREQKRKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLG
PFCACTSANNNTYWCMRTINETHNFLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQL
HVQLMELRSCKGYKQCNPRTRNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

## FIGURE 48

AACAAAGTTCAGTGACTGAGAGGGCTGAGCGGAGGCTGCTGAAGGGGGAGAAAGGAGTGAGGA  $\texttt{GCTGCTGGGCAGAGAGGGACTGTCCGGCTCCCAG} \underline{\textbf{ATG}} \texttt{CTGGGGCCTCCTGGGGAGCACAGCCC}$ TCGTGGGATGGATCACAGGTGCTGCTGTGGCGGTCCTGCTGCTGCTGCTGCTGCCACC TGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGGAACCGTACAGCTGCAGGGGGAAA CCGAGTCCGCCGGGCCCAGCCTTGGCCCTTCCGGCGGGGGCCACCTGGGAATCTTTCACC ATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCAC CCCCGCCACACCCTCACCACCACCACCACCACCCACCGCCACCACCATCCCCGCCA  $\mathsf{CGCTCGC}$ TTCCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGGTGAACGAGG GGAACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGCATGCCCCAGTGTACTATGGC AGCAGAGAATGGAGGAACACTGGGTCTGCAGTGCTGAAGGGTTTGGGGAGTGGAGAGCAAGG GTGCTCTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGACTCCCAGTGAGCCCCAGAAATG ACAAGCGTGTCTTGGCAGAGCCAGCACAAGTGGATGTGAAGTGCCCGTCTTGACCTCCTC ATCAGGCTGCTGCAGGCCTCTGGCGGGCAGGGCACTGGGAGAGGCCCTGAGAATGTCCTTTT GGTTTGGAGAAGGCAGTGTGAGGCTGCACAGTCAATTCATCGGTGCCTTAGTCCAAGAAAAT 

# FIGURE 49

 ${\tt MLGLLGSTALVGWITGAAVAVLLLLLLLATCLFHGRQDCDVERNRTAAGGNRVRRAQPWPFR} \\ {\tt RRGHLGIFHHRRHPGHVSHVPNVGLHHHHHPRHTPHHLHHHHHPRHAR} \\$ 

### FIGURE 50

 ${\tt GGCGGCTGCTGAGCTGCTTGAGGTGCAGTGTTGGGGGATCCAGAGCC} {\color{red} {\bf ATG} {\tt TCGGACCTGCTA} }$ CTACTGGGCCTGATTGGGGGCCTGACTCTTTACTGCTGACGCTGCTGGCCTTTGCCGG GTACTCAGGGCTACTGGCTGGGTGGAAGTGAGTGCTGGGTCACCCCCATCCGCAACGTCA CTGTGGCCTACAAGTTCCACATGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGC TGCAGCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCC CCCTGATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCC CTGAGCTCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGC TGTCCATCCTGCCTTGGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGG AGATCTACCAGGAAGACCAGATCCATTTCATGTGCCCACTGGCACGGCAGGGAGACTTCTAT GTGCCTGAGATGAAGGAGACAGAGTGGAAATGGCGGGGGCTTGTGGAGGCCATTGACACCCA GGTGGATGGCACAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCC CTGGCAGCCGGGAGACTTCAGCTGCCACACTGTCACCTGGGGCGAGCAGCCGTGGCTGGGAT GACGGTGACACCCGCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGA GGAGCTGGACTTGGAGGGCGAGGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGC ATGGCCTGCACCCTCCTGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCT CCAGCCCTCTTCCTCCTCTGGGGGAGGAGGGGTTCCTGAGGGACCTGACTTCCCCTGC TCCAGGCCTCTTGCTAAGCCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCA GGGACTATTTTCTGCACCAGCCCCCAGGGCTGCCGCCCCTGTTGTGTCTTTTTTTCAGACTC ACAGTGGAGCTTCCAGGACCCAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAAA AAAAAAAAA

## FIGURE 51

MSDLLLLGLIGGLTLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGR LFTESCSISPKLRSIAVYYDNPHMVPPDKCRCAVGSILSEGEESPSPELIDLYQKFGFKVFS FPAPSHVVTATFPYTTILSIWLATRRVHPALDTYIKERKLCAYPRLEIYQEDQIHFMCPLAR QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSAATLSPGAS SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEGPLGESRLDPGTEPLGTTKWLWEPTAPEK GKE

## FIGURE 52

# FIGURE 53

MTLRPSLLPLHLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFGDTLHI HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSHLAYGKRGF PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS KKKLKEEKRNKSKKK

## FIGURE 54

## FIGURE 55

### FIGURE 56

CTGCTGCATCCGGGTGTCTGGAGGCTGTGGCCGTTTTGTTTTCTTGGCTAAAATCGGGGGAG TGAGGCGGGCGCGCGCGCGCACACCGGGCTCCGGAACCACTGCACGACGGGGCTGGACTG ACCTGAAAAAA**ATG**TCTGGATTTCTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGG GAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGAT TATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCT GTGGTGTTATAGCAACCATAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGA GGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGG TTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTG CTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATCTTT  ${\tt TTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGACTTATGGCAG} {\tt TGA} {\tt ACACATCTGAT}$ TTGTAATGCCATTTTCTAAACTTATTTCTGAGTGTAGTCTCAGCTTAAAGTTGTGTAATACT AAAATCACGAGAACACCTAAACAACAACCAAAAATCTATTGTGGTATGCACTTGATTAACTT ATAAAATGTTAGAGGAAACTTTCACATGAATAATTTTTTGTCAAATTTTATCATGGTATAATT TGTAAAAATAAAAAGAAATTACAAAAGAAATTATGGATTTGTCAATGTAAGTATTTGTCATA TCTGAGGTCCAAAACCACAATGAAAGTGCTCTGAAGATTTAATGTGTTTATTCAAATGTGGT CTCTTCTGTGTCAAATGTTAAATGAAATATAAACATTTTTTAGTTTTTAAAATATTCCGTGG TCAAAATTCTTCCTCACTATAATTGGTATTTACTTTTACCAAAAATTCTGTGAACATGTAAT GTAACTGGCTTTTGAGGGTCTCCCAAGGGGTGAGTGGACGTGTTGGAAGAGAAGCACCAT GGTCCAGCCACCAGGCTCCCTGTGTCCCTTCCATGGGAAGGTCTTCCGCTGTGCCTCTCATT CCAAGGGCAGGAAGATGTGACTCAGCCATGACACGTGGTTCTGGTGGGATGCACAGTCACTC CACATCCACCACTG

# FIGURE 57

MSGFLEGLRCSECIDWGEKRNTIASIAAGVLFFTGWWIIIDAAVIYPTMKDFNHSYHACGVI ATIAFLMINAVSNGQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILFGGYVAKEK DIVYPGIAVFFQNAFIFFGGLVFKFGRTEDLWQ

## FIGURE 58

## FIGURE 59

TGGACGGACCTGAAAAAAATGTTTGGATTTNTAGAGGGNTTGAGATGTTCAGAATGCATGAC
TGGGGGAAAAGCGCAAATACTATTGCTTCCATTGCTGCTGGTGTANTATTTTTTTACAGGCTG
GTGGATTATCATAGATGCAGNTGTTATTTATCCCACCATGAAAGATTTCAACCANTCATACC
ATGCCTGTGGTGTTATAGCAACCATAGCCTTCNTAATGATTAATGCAGTATCGAATGGACAA
GTCCGAGGTGATAGTTACAGTGAAGGTTGTTTGGGTCAAACAGGTGCTCGCATTTGGCTTTT
CGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTT
ATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGNTGTATTTTTCCAGAATGCCTTC
ATCTTTTTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGANTTATGGCAGTG

## FIGURE 60

GGACACCGGGTTCCGGACCAATGCANGACGGGGTGGANTGACCTGAAAAAAATGTTTGGATT
TTTAGAGGGCTTGAGATGNTCAGAATGCATTGACTGGGGGAAAAGCGCAATANTATTGCTTT
CCATTGCTGCTGGTGTACTATTTTTTACAGGGTGGTGGATTATCATAGATGCAGCTGTTATT
TATCCCACCATGAAAGATTTNAACCACTCATACCATGCCTGTGGTGTTATAGCAACCATAGC
CTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTT
GTTTGGGTCAAACAGGTGNTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTTGGATTT
CTGATTGNATTCTATGCGGATTCTTCTTGGAGGTTATGTTGCTAAAGAAAAAAGACATAGTAT
ACCCTGGAATTNCTNTATTTTTCCAGAATGCC

## FIGURE 61

TAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATANTATTGCTTCC
ATTGNTGNTGGTGTANTATTTTTTTACAGGCTGGTGGATTATNATAGATGCAGCTGTTATTT
ATCCCACCATGAAAGATTTNAACCANTCATACCATGCCTGTGGTGTTATAGCAACCATAGCC
TTCCTAATGATTAATGCAGTATNGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTG
TTTGGGTCAAACAGGTGNTNGCATTTGGCTTTTNGTTGGTTTCATGTTGGCCTTTTGGATCTN
TGATTGCATTTATGTGGATTNTTTTTTGGAGGTTATGTTGCTAAAGNAAAAGACATAGTATAC
CCTGT

## FIGURE 62

### FIGURE 63

 ${\tt CGACGCCGGCGTG}$ TCTGCAAAGTTTACTTGGGACTATTCTCTGGCAGCTCCCCGAATCCTTTCTCCGAAGATGTCAAACGGCCCCCAGCGCCCCTGGTAACTGACAAGGAGGCCAGGAAGAAGGTTCTCAAACAAGC TTTTTCAGCCAACCAAGTGCCGGAGAAGCTGGATGTGGTGGTAATTGGCAGTGGCTTTGGGG GCCTGGCTGCAGCTGCAATTCTAGCTAAAGCTGGCAAGCGAGTCCTGGTGCTGGAACAACAT ACCAAGGCAGGGGGCTGCTGTCATACCTTTGGAAAGAATGGCCTTGAATTTGACACAGGAAT CCATTACATTGGGCGTATGGAAGAGGGCAGCATTGGCCGTTTTATCTTGGACCAGATCACTG AAGGGCAGCTGGACTGGGCTCCCCTGTCCTCTCTTTTGACATCATGGTACTGGAAGGGCCC AATGGCCGAAAGGAGTACCCCATGTACAGTGGAGAGAAAGCCTACATTCAGGGCCTCAAGGA GAAGTTTCCACAGGAGAAGCTATCATTGACAAGTATATAAAGCTTAAGGTGGTATCCA GTGGAGCCCCTCATGCCATCCTGTTGAAATTCCTCCCATTGCCCGTGGTTCAGCTCCTCGAC AGGTGTGGGCTGCTGACTCCTTTCTCCATTCCTTCAAGCATCCACCCAGAGCCTGGCTGA GGTCCTGCAGCAGCTGGGGGCCTCCTCTGAGCTCCAGGCAGTACTCAGCTACATCTTCCCCA CTTACGGTGTCACCCCCAACCACGGCCTTTTCCATGCACGCCCTGCTGGTCAACCACTAC ATGAAAGGAGGCTTTTATCCCCGAGGGGGTTCCAGTGAAATTGCCTTCCACACCATCCCTGT GATTCAGCGGGCTGGGGGCCTGTCCTCACAAAGGCCACTGTGCAGAGTGTGTTGCTGGACT CAGCTGGGAAAGCCTGTGGTGTCAGTGTGAAGAAGGGGCATGAGCTGGTGAACATCTATTGC CCCATCGTGGTCTCCAACGCAGGACTGTTCAACACCTATGAACACCTACTGCCGGGGAACGC GAGGAGTGGCAGGCGGAGCTGAAGGGAAAGCGGGGCAGTGACTATGAGACCTTCAAAAACTC CTTTGTGGAAGCCTCTATGTCAGTGGTCCTGAAACTGTTCCCACAGCTGGAGGGGAAGGTGG AGAGTGTGACTGCAGGATCCCCACTCACCAACCAGTTCTATCTGGCTGCTCCCCGAGGTGCC TGCTACGGGGCTGACCATGACCTGGGCCGCCTGCACCCTTGTGTGATGGCCTCCTTGAGGGC CCAGAGCCCCATCCCCAACCTCTATCTGACAGGCCAGGATATCTTCACCTGTGGACTGGTCG GGGCCCTGCAAGGTGCCCTGCTGCAGCAGCCCCATCCTGAAACCGGAACTTCATCTCAGAC CTTAAGAATCTTGATTCTAGGATCCGGGCACAGAAGAAAAAGAAT**TAG**TTCCATCAGGGAGG AGTCAGAGGAATTTGCCCAATGGCTGGGGCATCTCCCTTGACTTACCCATAATGTCTTTCTG
CATTAGTTCCTTGCACGTATAAAGCACTCTAATTTGGTTCTGATGCCTGAAGAGAGGCCTAG
TTTAAATCACAATTCCGAATCTGGGGCAATGGAATCACTGCTTCCAGCTGGGGCAGGTGAGA
TCTTTACGCCTTTTTATACACTGCCATCCCTACTAATAGGATATTTGACTTGGATAGCTTGATG TCTTTACGCCTTTTATAACATGCCATCCCTACTAATAGGATAITGACTTGGATAGCTIGATG
TCTCATGACGAGCGGGCGCTCTGCATCCCTACCCATGCCTCCTAACTCAGTGATCAAAGCGA
ATATTCCATCTGTGGATAGAACCCCTGGCAGTGTTGTCAGCTCAACCTGGTGGGTTCAGTTC
TGTCCTGAGGCTTCTGCTCTCATTCATTTAGTGCTACGCTGCACAGTTCTACACTGTCAAGG
GAAAAGGGAGACTAATGAGGCTTAACTCAAAACCTGGGCGTGGTTTTGGTTGCCATTCCATA
GGTTTGGAGAGCTCTAGATCTCTTTTTGTGCTGGGTTCAGTGGCTCTTCAGGGGACAGGAAAT
GCCTGTGTCTGGCCAGTGTGGTTCTGGAGCTTTTGGGGGTAACAGCAGGATCCATCAGTTAGTA CTTATCCACCAAATACACAGGGAAGGGTGATGCAGGGAAGGGTGACATCAGGAGTCAGGGCA TGGACTGGTAAGATGAATACTTTGCTGGGCTGAAGCAGGCTGCAGGGCATTCCAGCCAAGGGCACAGCAGGGGACAGTGCAGGGGAGGTGTGGGGTAAGGGAAGTCACATCAGAAAAGGGA AAGCCACGGAATGTGTGAAGCCCAGAAATGGCATTTGCAGTTAATTAGCACATGTGAGGG TTAGACAGGTAGGTGAATGCAAGCTCAAGGTTTGGAAAAATGACTTTTCAGTTATGTCTTTG GTATCAGACATACGAAAGGTCTCTTTGTAGTTCGTGTTAATGTAACATTAATAAATTTATTG ATTCCATTGCTTTAAAAAAAAAAAAAAAAAAA

## FIGURE 64

MWLPLVLLLAVLLLAVLCKVYLGLFSGSSPNPFSEDVKRPPAPLVTDKEARKKVLKQAFSAN QVPEKLDVVVIGSGFGGLAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLEFDTGIHYIG RMEEGSIGRFILDQITEGQLDWAPLSSPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQ EEAIIDKYIKLVKVVSSGAPHAILLKFLPLPVVQLLDRCGLLTRFSPFLQASTQSLAEVLQQ LGASSELQAVLSYIFPTYGVTPNHSAFSMHALLVNHYMKGGFYPRGGSSEIAFHTIPVIQRA GGAVLTKATVQSVLLDSAGKACGVSVKKGHELVNIYCPIVVSNAGLFNTYEHLLPGNARCLP GVKQQLGTVRPGLGMTSVFICLRGTKEDLHLPSTNYYVYYDTDMDQAMERYVSMPREEAAEH IPLLFFAFPSAKDPTWEDRFPGRSTMIMLIPTAYEWFEEWQAELKGKRGSDYETFKNSFVEA SMSVVLKLFPQLEGKVESVTAGSPLTNQFYLAAPRGACYGADHDLGRLHPCVMASLRAQSPI PNLYLTGQDIFTCGLVGALQGALLCSSAILKRNLYSDLKNLDSRIRAQKKKN

### FIGURE 65

GCAGCGGCGAGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCTA  $\mathsf{GGGGTTGGCACCGGCCCCGAGAGGAGGAGG}$   $\mathsf{CGGGTCCGGATAGGGCTGACGCTGCTGTG}$ TGCGGTGCTGCTGAGCTTGGCCTCGGCGTCCTCGGATGAAGAAGGCAGCCAGGATGAATCCT GTAGTTGCTGGTCAAATATTTCTTGATTCAGAAGAATCTGAATTAGAATCCTCTATTCAAGA AGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAAAGTGTCACAGAAGATATCAGCTTTCTAG ACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCTTTTCCTAGATAA GGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGCTACAACCT ATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAAACTGAAGAAGAGGCTGCTAAGAGA CGGCAGATGCAGGAAGCAGAAATGATGTATCAAACTGGAATGAAAATCCTTAATGGAAGCAA CCAAAGCCCTGGAGAGAGTGTCATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATC CAGGCAGCGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGC TCTTGGCTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAAGTCAGGCAAAGGCTCTTGTAT ATTATACATTTGGAGCTCTTGGGGGCAATCTAATAGCCCACATGGTTTTGGTAAGTAGACTT **TAG**TGGAAGGCTAATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACTTTT ATTCTTGTTAATGGATATAACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACA ATTTTTCTTTAAAATGATTAGTTTGGCTGATTGCCCCTAAAAAGAGAGATCTGATAAATGGC TCTTTTTAAATTTTCTCTGAGTTGGAATTGTCAGAATCATTTTTTACATTAGATTATCATAA TTTTAAAAATTTTTCTTTAGTTTTTCAAAATTTTTGTAAATGGTGGCTATAGAAAAACAACAT GAAATATTATACAATATTTTGCAACAATGCCCTAAGAATTGTTAAAAATTCATGGAGTTATTT GTGCAGAATGACTCCAGAGAGCTCTACTTTCTGTTTTTTACTTTTCATGATTGGCTGTCTTC CCATTTATTCTGGTCATTTATTGCTAGTGACACTGTGCCTGCTTCCAGTAGTCTCATTTTCC CTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGGAAGATTAACTCATTTTTAATAAA 

## FIGURE 66

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLDSKTTLTSDESVKDHTTAGRVVAGQIFLD SEESELESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHG EPCHFPFLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAAKRRQMQEAEMM YQTGMKILNGSNKKSQKREAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFEK LTEEGSPKGQTALGFLYASGLGVNSSQAKALVYYTFGALGGNLIAHMVLVSRL

## FIGURE 67

# FIGURE 68

 ${\tt MACRCLSFLLMGTFLSVSQTVLAQLDALLVFPGQVAQLSCTLSPQHVTIRDYGVSWYQQRAGSAPRYLLYYRSEEDHHRPADIPDRFSAAKDEAHNACVLTISPVQPEDDADYYCSVGYGFSP}$ 

### FIGURE 69

CCCGCCGAGGTCCGGACAGGCCGAG<mark>ATG</mark>ACGCCGAGCCCCTGTTGCTGCTCCTGCTGCCGC CGCTGCTGCTGGGGGCCTTCCCACCGGCCGCCGCCGCCGAGGCCCCCAAAGATGGCGGACAAGGTGGTCCCACGGCAGGTGGCCCGCTGGGCCGCACTGTGCGGCTGCAGTGCCCAGTGGA GGGGGACCCGCCGCCGCTGACCATGTGGACCAAGGATGGCCGCACCATCCACAGCGGCTGGA GCCGCTTCCGCGTGCTGCCGCAGGGGCTGAAGGTGAAGCAGGTGGAGCGGGAGGATGCCGGCGTGTACGTGTGCAAGGCCACCAACGGCTTCGGCAGCCTGAGCGTCAACTACACCCTCGTCGT GCTGGATGACATTAGCCCAGGGAAGGAGCCTGGGGCCCGACAGCTCCTCTGGGGGTCAAG AGGACCCCGCCAGCCAGCAGTGGGCACGACCGCGCTTCACACAGCCCTCCAAGATGAGGCGC CGGGTGATCGCAGCAGCAGTGGGTAGCTCCGTGCGCTCAAGTGCGTGGCCAGCGGCACCC TCGCCCGACATCACGTGGATGAAGGACGACCAGGCCTTGACGCGCCCAGAGGCCGCTGAGC CCAGGAAGAAGAAGTGGACACTGAGCCTGAAGAACCTGCGGCCGGAGGACAGCGGCAAATAC ACCTGCCGCGTGTCGAACCGCGCGGGCGCCATCAACGCCACCTACAAGGTGGATGTGATCCA GCGGACCCGTTCCAAGCCCGTGCTCACAGGCACGCACCCCGTGAACACGACGGTGGACTTCG GGGGGACCACGTCCTTCCAGTGCAAGGTGCGCAGCGACGTGAAGCCGGTGATCCAGTGGCTG AAGCGCGTGGAGTACGGCGCCGAGGGCCCACAACTCCACCATCGATGTGGGCGGCCAGAA GTTTGTGGTGCTGCCCACGGGTGACGTGTGGTCGCGGCCCGACGGCTCCTACCTCAATAAGC CAGCGGAGACAAGGACCTTCCCTCGTTGGCCGCCCTCAGCGCTGGCCCTGGTGTGGGGCTGT ACACTCACACGTGGAGGGCAAGGTCCACCAGCACTATCAGTGC**TAG**ACGGCACCGT ATCTGCAGTGGGCACGGGGGGGCCGGCCAGACAGGCAGACTGGGAGGATGGAGGACGGAGCT GCAGACGAAGGCAGGGGACCCATGGCGAGGAGGAATGGCCAGCACCCCAGGCAGTCTGTGTG CACACACACACGGATATGCTGTCTGGACGCACACACGTGCAGATATGGTATCCGGACACA GCCTGGACACACGCAGATATGCTGTCTAGTCACACACACGCAGACATGCTGTCCGGACAC TATTGCCTGGACACACACATGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATA TGCTGTCCGGATACACACGCACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGA CACACATGCACACAGGTGCAGATATGCTGCCTGGACACACGCGGGCACACACTTTTGG GAGGGTGTGCCGTGAAGCCTGCAGTACGTGTGCCGTGAGGCTCATAGTTGATGAGGGACTTT CCCTGCTCCACCGCCCCAACTCTGCCCGCCTCTGTCCCCGCCTCAGTCCCCGCCTC CATCCCCGCCTCTGTCCCCTGGCCTTGGCGGCTATTTTTGCCACCTGCCTTGGGTGCCCAGG AGTCCCTACTGCTGTGGGCTGGGGTTGGGGGCACAGCCCCAAGCCTGAGAGGCTGGAGCCCATGGCTAGTGCTATCCCCAGTGCATTCTCCCCCTGACACAGAGAAGGGGCCTTGGTA TTTATATTTAAGAAATGAAGATAATATTAATAATGATGGAAGGAAGACTGGGTTGCAGGGAC TGTGGTCTCTCCTGGGGCCCGGGACCCGCCTGGTCTTTCAGCCATGCTGATGACCACACCCC GTCCAGGCCAGACACCACCCCCACCCCACTGTCGTGGTGGCCCCAGATCTCTGTAATTTTA 

## FIGURE 70

MTPSPLLLLLPPLLLGAFPPAAAARGPPKMADKVVPRQVARLGRTVRLQCPVEGDPPPLTM
WTKDGRTIHSGWSRFRVLPQGLKVKQVEREDAGVYVCKATNGFGSLSVNYTLVVLDDISPGK
ESLGPDSSSGGQEDPASQQWARPRFTQPSKMRRRVIARPVGSSVRLKCVASGHPRPDITWMK
DDQALTRPEAAEPRKKKWTLSLKNLRPEDSGKYTCRVSNRAGAINATYKVDVIQRTRSKPVL
TGTHPVNTTVDFGGTTSFQCKVRSDVKPVIQWLKRVEYGAEGRHNSTIDVGGQKFVVLPTGD
VWSRPDGSYLNKLLITRARQDDAGMYICLGANTMGYSFRSAFLTVLPDPKPPGPPVASSSSA
TSLPWPVVIGIPAGAVFILGTLLLWLCQAQKKPCTPAPAPPLPGHRPPGTARDRSGDKDLPS
LAALSAGPGVGLCEEHGSPAAPQHLLGPGPVAGPKLYPKLYTDIHTHTHTHSHTHSHVEGKV
HQHIHYQC

### FIGURE 71

## FIGURE 72

MVGTKAWVFSFLVLEVTSVLGRQTMLTQSVRRVQPGKKNPSIFAKPADTLESPGEWTTWFNI DYPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQ RPGQNCSNYTVRFLCPPGSLRRDTERIWSPWSPWSKCSAACGQTGVQTRTRICLAEMVSLCS EASEEGQHCMGQDCTACDLTCPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAIYLLTK TPKLLTQTDSDGRFRIPGLCPDGKSILKITKVKFAPIVLTMPKTSLKAATIKAEFVRAETPY MVMNPETKARRAGQSVSLCCKATGKPRPDKYFWYHNDTLLDPSLYKHESKLVLRKLQQHQAG EYFCKAQSDAGAVKSKVAQLIVTASDETPCNPVPESYLIRLPHDCFQNATNSFYYDVGRCPV KTCAGQQDNGIRCRDAVQNCCGISKTEEREIQCSGYTLPTKVAKECSCQRCTETRSIVRGRV SAADNGEPMRFGHVYMGNSRVSMTGYKGTFTLHVPQDTERLVLTFVDRLQKFVNTTKVLPFN KKGSAVFHEIKMLRRKEPITLEAMETNIIPLGEVVGEDPMAELEIPSRSFYRQNGEPYIGKV KASVTFLDPRNISTATAAQTDLNFINDEGDTFPLRTYGMFSVDFRDEVTSEPLNAGKVKVHL DSTQVKMPEHISTVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLFNL DVPESRRCFVKVRAYRSERFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGA CVPAFCDDQSPDAYSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLNYRRTDHEDPR VKKTAFQISMAKPRPNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYNTVPFN EDDPMSWTEDYLAWWPKPMEFRACYIKVKIVGPLEVNVRSRNMGGTHRRTVGKLYGIRDVRS TRDRDQPNVSAACLEFKCSGMLYDQDRVDRTLVKVIPQGSCRRASVNPMLHEYLVNHLPLAV NNDTSEYTMLAPLDPLGHNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALT FNCVERQVGRQSAFQYLQSTPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFPRVA OOPLIN

## FIGURE 73

ACCTACCTACCCGTACGCATACATACATATGTGTATATATGTAAACTAGACAAAGATCGCAGATCATAAAGC AAGCTCTGCTTTAGTTTCCAAGAAGATTACAAAGAATTTAGAG ATG TATTTGTCAAGATCCCTGTCGATTCATG $\hbox{\tt CCCTTTGGGTTACGGTGTCCTCAGTGATGCAGCCCTACCCTTTGGTTTGGGGACATTATGATTTGTGTAAGACT}$ CAGATTTACACGGAAGAAGGGAAAGTTTGGGATTACATGGCCTGCCAGCCGGAATCCACGGACATGACAAAATA TCTGAAAGTGAAACTCGATCCTCCGGATATTACCTGTGGAGACCCTCCTGAGACGTTCTGTGCAATGGGCAATC  $\tt CCTACATGTGCAATAATGAGTGTGATGCGAGTACCCCTGAGCTGGCACACCCCCTGAGCTGATGTTTGATTTT$ GAAGGAAGACATCCCTCCACATTTTGGCAGTCTGCCACTTGGAAGGAGTATCCCAAGCCTCTCCAGGTTAACAT CACTCTGTCTTGGAGCAAAACCATTGAGCTAACAGACAACATAGTTATTACCTTTGAATCTGGGCGTCCAGACC AAATGATCCTGGAGAAGTCTCTCGATTATGGACGAACATGGCAGCCCTATCAGTATTATGCCACAGACTGCTTA GATGCTTTTCACATGGATCCTAAATCCGTGAAGGATTTATCACAGCATACGGTCTTAGAAATCATTTGCACAGA AGAGTACTCAACAGGGTATACAACAAATAGCAAAATAATCCACTTTGAAATCAAAGACAGGTTCGCGCTTTTTG  $\tt CTGGACCTCGCCTACGCAATATGGCTTCCCTCTACGGACAGCTGGATACAACCAAGAAACTCAGAGATTTCTTT$ ACAGTCACAGACCTGAGGATAAGGCTGTTAAGACCAGCCGTTGGGGAAATATTTGTAGATGAGCTACACTTGGC ACGCTACTTTTACGCGATCTCAGACATAAAGGTGCGAGGAAGGTGCAAGTGTAATCTCCATGCCACTGTATGTG TGTATGACAACAGCAAATTGACATGCGAATGTGAGCACAACACTACAGGTCCAGACTGTGGGAAATGCAAGAAG AATTATCAGGGCCGACCTTGGAGTCCAGGCTCCTATCTCCCCATCCCCAAAGGCACTGCAAATACCTGTATCCC CAGTATTTCCAGTATTGGTACGAATGTCTGCGACAACGAGCTCCTGCACTGCCAGAACGGAGGGACGTGCCACA  $\tt GGGAACCGCCAGCCCCTGGTGTTC{\color{red}{\textbf{TAG}}} GTGTCACCTCCAGCCACACCGGACGGCCTGTGCCGTGGGGAAGCA$ CTAAGAAGGCCTAACTGAACTAAGCCATATTTATCACCCGTGGACAGCACATCCGAGTCAAGACTGTTAATTTC TGACTCCAGAGGAGTTGGCAGCTGTTGATATTATCACTGCAAATCACATTGCCAGCTGCAGAGCATATTGTGGA TACTCTAGCGTGGTGCGCCCTAGTACGACTCCGCCCAGTGTCTGGACCAAACCAAATAGCATTCTTTGCTGTCAG GTGCATTGTGGGCATAAGGAAATCTGTTACAAGCTGCCATATTGGCCTGCTTCCGTCCCTGAATCCCTTCCAAC TGTGTAACAGCCCCCTCTAAAAGCGCAAGCCAGTCATACCCCTGTATATCTTAGCAGCACTGAGTCCAGTGCGA ATTTTTCTTGAACTACTGTAATATGTAGATTTTTTGTATTATTGCCAATTTGTGTTACCAGACAATCTGTTAAT GATTTCTCTGTAAGGGCAACGAACGTGCTGGCATCAAAGAATATCAGTTTACATATAACAAGTGTAATAAGA TTCCACCAAAGGACATTCTAAATGTTTTCTTGTTGCTTTAACACTGGAAGATTTAAAGAATAAAAACTCCTGCA TTACTGATTTCTGTGTGGACTGAGTACATTCAGCTGACGAATTTAGTTCCCCAGGAAGATGGATTGATGTTCACT AAAAAA

# FIGURE 74

MYLSRSLSIHALWVTVSSVMQPYPLVWGHYDLCKTQIYTEEGKVWDYMACQPESTDMTKYLK
VKLDPPDITCGDPPETFCAMGNPYMCNNECDASTPELAHPPELMFDFEGRHPSTFWQSATWK
EYPKPLQVNITLSWSKTIELTDNIVITFESGRPDQMILEKSLDYGRTWQPYQYYATDCLDAF
HMDPKSVKDLSQHTVLEIICTEEYSTGYTTNSKIIHFEIKDRFALFAGPRLRNMASLYGQLD
TTKKLRDFFTVTDLRIRLLRPAVGEIFVDELHLARYFYAISDIKVRGRCKCNLHATVCVYDN
SKLTCECEHNTTGPDCGKCKKNYQGRPWSPGSYLPIPKGTANTCIPSISSIGTNVCDNELLH
CQNGGTCHNNVRCLCPAAYTGILCEKLRCEEAGSCGSDSGQGAPPHGTPALLLLTTLLGTAS
PLVF

## FIGURE 75

CCCACGCGTCCGGGTGACCTGGGCCGAGCCCTCCCGGTCGGCTAAGATTGCTGAGGAGGCGG CGGGTAGCTGGCAGGCCGACTTCCGAAGGCCGCCGTCCGGGCGAGGTGTCCTCATGACTT  $\tt CTCTTGTGGACC{\color{red} ATG}{\color{blue} TCCGTGATCTTTTTTGCCTGCGTGGTACGGGTAAGGGATGGACTGCC}$ CCTCTCAGCCTCTACTGATTTTTACCACACCCAAGATTTTTTGGAATGGAGGAGACGGCTCA AGAGTTTAGCCTTGCGACTGGCCCAGTATCCAGGTCGAGGTTCTGCAGAAGGTTGTGACTTT AGTATACATTTTTCTTCTTCGGGGACGTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCC AGCAGCCATGGCCTTCTGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCCTATGACA CTACCTGCATTGGCCTAGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTCAG AAAGTGAAGTGGCATTTTAACTATGTAAGTTCCTCTCAGATGGAGTGCAGCTTGGAAAAAAT TCAGGAGGAGCTCAAGTTGCAGCCTCCAGCGGTTCTCACTCTGGAGGACACAGATGTGGCAA ATGGGGTGATGAATGGTCACACACCGATGCACTTGGAGCCTGCTCCTAATTTCCGAATGGAA CCAGTGACAGCCCTGGGTATCCTCTCCCTCATTCTCAACATCATGTGTGCTGCCCTGAATCT CATTCGAGGAGTTCACCTTGCAGAACATTCTTTACAGGATCCAAGGAGCTGGTTCTGCTGGT  $\texttt{TGGACCAAACCTCG} \underline{\textbf{TGA}} \texttt{GCCAGCCACCCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCAT}$ CCGGGAGCAGTGATGTCAAACTTCTGCTGCTGGGGAAATCTCATCAGCAGGGAGCCTGTGGA AAAGGGCATGTCAGTGAAATCTGGGAATGGCTGGATTCGGAAACATCTGCCCATGTGTATTG ATGGCAGAGCTGTTGCCCACAAGCGCCTTTTATTTAGGGTAAAATTAACAAATCCATTCTAT TCCTCTGACCCATGCTTAGTACATATGACCTTTAACCCTTACATTTATATGATTCTGGGGTT GCTTCAGAAGTGTTATTTCATGAATCATTCATATGATTTGATCCCCCAGGATTCTATTTTGT TTAATGGGCTTTTCTACTAAAAGCATAAAATACTGAGGCTGATTTAGTCAGGGCAAAACCAT TTACTTTACATATTCGTTTTCAATACTTGCTGTTCATGTTACACAAGCTTCTTACGGTTTTC TTGTAACAATAAATATTTTGAGTAAATAATGGGTACATTTTAACAAACTCAGTAGTACAACC TAAACTTGTATAAAAGTGTGTAAAAATGTATAGCCATTTATATCCTATGTATAAATTAAATG AAAAG

# FIGURE 76

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFSIHF SSFGDVACMAICSCQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQKVKW HFNYVSSSQMECSLEKIQEELKLQPPAVLTLEDTDVANGVMNGHTPMHLEPAPNFRMEPVTA LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

# FIGURE 77

 ${\tt TGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCNTATGACACTACCTGCATTGGCNT} \\ AGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTCAGAAAGTGAAGTGCATT\\ {\tt TTAACTATGTAAGTTCCTNTCAGATGGAGTGCAGCTTGGAAAAAATTCAGGAGGAGCTCAAG} \\ {\tt TTGCAGCCTCCAGCGGTTCTCANTATGGAGGACACAGATGTGGCAAATGGGGT} \\ \\$ 

## FIGURE 78

CTCAGCGGCGCTTCCTCGTAGCGAGCCTAGTGGCGGGTGTTTGCATTGAAACGTGAGCGCGA CCCGACCTTAAAGAGTGGGGAGCAAAGGGAGGACAGAGCCCTTTAAAAACGAGGCGGGTGGTG CCTGCCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTT TCTGTCGCAGGCTGCGAGGAAAGGCCCCTAGGCTGGGTCTGGGTGCTTGGCGGCGGCGCTT  $\tt CCTCCCGGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGT\underline{\textbf{A}}$ TG GAAGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGTGTATTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGAC CCGCTTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGA TTGCGCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCCTGCTCCTGCCC TTCTCCATCATCAGCAATGAGGTGCTGCTCTCCCTGCCTCGGAACTACTACATCCAGTGGCT CAACGGCTCCCTCATCCATGGCCTCTGGAACCTTGTTTTTCTCTTCCCCCAACCTGTCCCTCA TCTTCCTCATGCCCTTTGCATATTTCTTCACTGAGTCTGAGGGCTTTGCTGGCTCCAGAAAG GGTGTCCTGGGCCGGGTCTATGAGACAGTGGTGATGTTGATGCTCCTCACTCTGCTGGTGCT AGGTATGGTGTGGGCATCAGCCATTGTGGACAAGAACAAGGCCAACAGAGAGTCACTCT CTGCTCCTGGTGTACTCCACTGGGTCTCGCCCGCATGTTCTCCGTCACTGGGAAGCTGCT AGTCAAGCCCCGGCTGCTGGAAGACCTGGAGGAGCAGCTGTACTGCTCAGCCTTTGAGGAGG CTACACAGACAGGTCCTGGCTCTGCAGACACAGAGGGTCCTGCTGGAGAAGAGGCGGAAGGC GCCTGTCTGTGCTCATTGTGGCCATCCACATCCTGGAGCTGCTCATCGATGAGGCTGCCATG CCCCGAGGCATGCAGGGTACCTCCTTAGGCCAGGTCTCCTTCTCCAAGCTGGGCTCCTTTGG  ${ t TGCCGTCATTCAGGTTGTACTCATCTTTTACCTAATGGTGTCCTCAGTTGTGGGCCTTCTATA}$ GCTCTCCACTCTTCCGGAGCCTGCGGCCCAGATGGCACGACACTGCCATGACGCAGATAATT GGGAACTGTGTCTGTCTCCTGGTCCTAAGCTCAGCACTTCCTGTCTTCTCTCGAACCCTGGG ACTGCAGCTGTGCGGGCAGAGCTGATCCGGGCCTTTGGGCTGGACAGACTGCCGCTGCCCGT  $\tt CTCCGGTTTCCCCCAGGCATCTAGGAAGACCCAGCACCAG{\color{blue} {\bf TGA}} CCTCCAGCTGGGGGTGGGA$ AGGAAAAAACTGGACACTGCCATCTGCTGCCTAGGCCTGGAGGGAAGCCCAAGGCTACTTGG ACCTCAGGACCTGGAATCTGAGAGGGTGGGTGGCAGAGGGGAGCAGAGCCATCTGCACTATT GCATAATCTGAGCCAGAGTTTGGGACCAGGACCTCCTGCTTTTCCATACTTAACTGTGGCCT CAGCATGGGGTAGGGCTGGGTGACTGGGTCTAGCCCCTGATCCCAAATCTGTTTACACATCA ATCTGCCTCACTGCTGTTCTGGGCCATCCCCATAGCCATGTTTACATGATTTGATGTGCAAT CTTGCCTCTGGCCCAGCAGAGCCTAAGCACTGTGCTATCCTGGAGGGGCTTTGGACCACCTG AAAGACCAAGGGGATAGGGAGGAGGAGGCTTCAGCCATCAGCAATAAAGTTGATCCCAGGGA AAAAAA

# FIGURE 79

MEAPDYEVLSVREQLFHERIRECIISTLLFATLYILCHIFLTRFKKPAEFTTVDDEDATVNK
IALELCTFTLAIALGAVLLLPFSIISNEVLLSLPRNYYIQWLNGSLIHGLWNLVFLFPNLSL
IFLMPFAYFFTESEGFAGSRKGVLGRVYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESL
YDFWEYYLPYLYSCISFLGVLLLLVCTPLGLARMFSVTGKLLVKPRLLEDLEEQLYCSAFEE
AALTRRICNPTSCWLPLDMELLHRQVLALQTQRVLLEKRRKASAWQRNLGYPLAMLCLLVLT
GLSVLIVAIHILELLIDEAAMPRGMQGTSLGQVSFSKLGSFGAVIQVVLIFYLMVSSVVGFY
SSPLFRSLRPRWHDTAMTQIIGNCVCLLVLSSALPVFSRTLGLTRFDLLGDFGRFNWLGNFY
IVFLYNAAFAGLTTLCLVKTFTAAVRAELIRAFGLDRLPLPVSGFPQASRKTQHQ

# FIGURE 80

GGCTGCCGAGGGAAGGCCCCTTGGGTTGGTCTTGGTTGCTTGGCGGCGGCGGTTTCNTCCCC
GCTCGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGAAGC
ACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGTGTA
TTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGACCCGCTTC
AAGAAGCCTGCTGAGTTCACCACAGTGGATGAAGATGCCACCG

## FIGURE 81

## FIGURE 82

# FIGURE 83

 $\verb|MLLWVILLVLAPVSGQFARTPRPIIFLQPPWTTVFQGERVTLTCKGFRFYSPQKTKWYHRYLGKEILRETPDNILEVQESGEYRCQAQGSPLSSPVHLDFSSEMGFPHAAQANVELLGSSDLLT$ 

## FIGURE 84

CAGAAGAGGGGGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCGGTGT GAGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGCGGAGGAGGCTGTGAG GAGTGTGTGGAACAGGACCCGGGACAGAGGAACCATGGCTCCGCAGAACCTGAGCACCTTTT GCCTGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGGACGAGATTTCTATAAGATCTTG GGGGTGCCTCGAAGTGCCTCTATAAAGGATATTAAAAAGGCCTATAGGAAACTAGCCCTGCA GCTTCATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCCAGGATCTGGGTG CTGCTTATGAGGTTCTGTCAGATAGTGAGAAACGGAAACAGTACGATACTTATGGTGAAGAA GGATTAAAAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTCACACTTCTTTGGGGATTT TGGTTTCATGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTCCAAGAGGAAGTGATA TTATTGTAGATCTAGAAGTCACTTTGGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTT AGAAACAAACCTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTCGGCAAGAGAT GCGGACCACCCAGCTGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAAT GCCCTAATGTCAAACTAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTG AGAGACGGCATGGAGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGG TGTACACAAATGTGACAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACT CACTTGGATGGTCACAAGGTACATATTTCCCGGGATAAGATCACCAGGCCAGGAGCGAAGCT ATGGAAGAAAGGGGAAGGGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAA TCACTTTTGATGTGGATTTTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAA CAGCTACTGAAACAAGGGTCAGTGCAGAAGGTATACAATGGACTGCAAGGATAT**TGA**GAGTG TCATCATGAAATGAATAAGAGGGCTTAAGAATTTGTCCATTTGCATTCGGAAAAGAATGACC AGCAAAAGGTTTACTAATACCTCTCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGT TTCAAGAATTAAAGCTGCAAGAGGACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGA GTTGTTAGCAATTTCATTCAAAATGCCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTTG TTATTTTTA

## FIGURE 85

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQ
AQEKFQDLGAAYEVLSDSEKRKQYDTYGEEGLKDGHQSSHGDIFSHFFGDFGFMFGGTPRQQ
DRNIPRGSDIIVDLEVTLEEVYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQ
MTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGMEYPFIGEGEPHVDGEPGDLRFRIKVVKH
PIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAKLWKKGEGLPNFD
NNNIKGSLIITFDVDFPKEOLTEEAREGIKQLLKQGSVOKVYNGLQGY

#### Important features:

### Signal peptide:

amino acids 1-22

#### Cell attachment sequence.

amino acids 254-257

### Nt-dnaJ domain signature.

amino acids 67-87

### Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

### N-glycosylation site.

amino acids 5-9, 261-265

### Tyrosine kinase phosphorylation site.

amino acids 253-260

### N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

#### Amidation site.

amino acids 164-168

## FIGURE 86

TGGGACCAGGGAACCCCGGGCCCCCCGGTGGAGNGCCTAACAGGCCGGTGGNTGCGACCGAA
GCGGCGGGCGGAGGAGGTTTTGAGGATTTTTGGAACAGGACCCGGACAGAGGAACCATGGTT
CCGCAGAACNTGAGCACNTTTTGCCTGTTGNTGNTATACTTCATCGGGGCGGTGATTGCCGG
ACGAGATTTNTATAAGATTTTGGGGTGCCTNGAAGTGCCTTNTATAAAGGATATTAAAAAGG
CCTATAGGAAACTAGCCCTGCAGNTTTATCCCGACCGGAACCCTGATGATCCACAAGCCCAG
GAGAAATTCCAGGATTTGGGTGCTGCTTATGAGGTTNTGTCAGATAGTGAGAAACGGAAACA
GTACGATAATTATGGTGAAGAAGGATTAAAAAGATGGTNATCAGAGCTCCCATGGAGACATTT
TTTCACACTTNTTTGGGGATTTTGGTTTCATGTTTTGGAGGAACCCCTNGTCAGCAAGACAGA
AATATTCCAAGAG

## FIGURE 87

GGCACGAGGCGGGGGCAGTCGCGGGATGCGCCCGGGAGCCACAGCCTGAGGCCCTCAGGT CTCTGCAGGTGTCGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCA AGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTTCTGGCTTCGTTTGCAGCC TTGGTGCTGGTTTGCAGGCAGCGCTACTGCCGGCCGAGACCTGCTGCAGCGCTATGATTC TAAGCCCATTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC TGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCCATTCTGGAGAATGAAGACTGGATC GAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGAC AGAGAAGCTTGTTGCCATGACAATGGGCTCTGGGGCCAAGATGAAGACTTCAGCCAGTGTCA GCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGGATGATGTTGTGAAGTCGATG TCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACC AGTCTCTGTCGGCTGCTGAGGAGCATTTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAG  $\tt CCAGATAAAGGCCTCCCAGGCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATT{\color{red}{\textbf{TAG}}} TGCCT$ ACAGGCCAGCAGCTAGCCATGAAGGCCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTT TAAAGCAGGAGATCCCCGTCAGTTTATGCCTCTTTTTGCAGTTGCAAACTGTGGCTGGTGAGT GGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACTCTCTGCAAGAGGAGTATTGAAAA CTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCT AAGAAATCAAGAGGTTTCACATTAAAATTAGAATTTCTGGCCTCTCTCGATCGGTCAGAATG GGTCCCTGAGGCGTCTGGGGTCTCTCCTCTCCCTTGCAGGTTTGGGGTTTGAAGCTGAGGAACT ACAAAGTTGATGATTTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTG 

# FIGURE 88

METVVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPSEL ELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKTSAS VSDIIVVAKRISPRVDDVVKSMYPPLDPKLLDARTTALLLSVSHLVLVTRNACHLTGGLDWI DQSLSAAEEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

## FIGURE 89

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCC CCACCACTGCAGCCATGATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGA TTTGGAGTGTTTTTCCTGTTCTTTGGAATGATTCTCTTTTTTTGACAAAGCACTACTGGCTAT CTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTTCTCTTGTTCAG GGGCTTCTTTCCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCTTGGATCCCTCCTAAAT TTACCTGGAATTAGATCATTTGTAGATAAAGTTGGAGAAAGCAACAATATGGTA**TAA**CAACA GCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTTTACAGGAGTTTAAAACGTATAG CCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTACTCAAGTGA ACTAAGAAGAAGTCAGCAAGCAAACTGAGAGAGGTGAAATCCATGTTAATGATGCTTAAGAA ACTCTTGAAGGCTATTTGTGTTTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAA CTGTGGTGCCTGTTTCTTTTTTTTTTTTTTTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCT TTTTAGAAGTGTCCACTGCAATGGCAAAAATATTTCCAGTTGCACTGTATCTCTGGAAGTGA TGCATGAATTCGATTGGATTGTGTCATTTTAAAGTATTAAAACCAAGGAAACCCCAATTTTG ATGTATGGATTACTTTTTTTTTGNGCNCAGGGCC

# FIGURE 90

MISLTDTQKIGMGLTGFGVFFLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQK HKMKATGFFLGGVFVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFIRRVPVLGSLLNLPGI RSFVDKVGESNNMV

### Important features:

### Transmembrane domains:

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

### N-myristoylation sites.

amino acids 11-16, 51-56 and 116-121

# Aminoacyl-transfer RNA synthetases class-II protein.

amino acids 49-59

# FIGURE 91

# FIGURE 92

GGCACGAGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAA CTGTGGGTAGTTATTTATTTCTGAATAAGAGCGTCCACGCATC**ATG**GACCTCGCGGGACTGC TGAAGTCTCAGTTCCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGCTAATC ATCAACACCATTCAGCTCTTCACTCTCCTCCTCGGCCCATTAACAAGCAGCTCTTCCGGAA GATCAACTGCAGACTGTCCTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGT GCCATCGTGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGCTGGAGCCTGTCCGA TTATCGGCTGGATGTGGTACTTCACCGAGATGGTCTTCTGTTCGCGCAAGTGGGAGCAGGAT CGCAAGACGGTTGCCACCAGTTTGCAGCACCTCCGGGACTACCCCGAGAAGTATTTTTTCCT GATTCACTGTGAGGGCACACGGTTCACGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCC GGGCCAAGGGGCTGCCTCGCCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATC ACCGTGAGGAGCTTGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTTCAGAAA TTAGGAGGATCCCACTGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCAC AAGCTCTACCAGGAGAAGGATGCCTTTCAGGAGGAGTACTACAGGACGGGCACCTTCCCAGA GACGCCCATGGTGCCCCCCGGCGGCCCTGGACCCTCGTGAACTGGCTGTTTTGGGCCTCGC TGGTGCTCTACCCTTTCTTCCAGTTCCTGGTCAGCATGATCAGGAGCGGGTCTTCCCTGACG **GA**CTCAGGGAGGTGTCACCATCCGAAGGGAACCTTGGGGAACTGGTGGCCTCTGCATATCCT CCTTAGTGGGACACGGTGACAAAGGCTGGGTGAGCCCCTGCTGGGCACGGCGGAAGTCACGA CCTCTCCAGCCAGGGAGTCTGGTCTCAAGGCCGGATGGGGAGGAAGATGTTTTGTAATCTTT TTTTCCCCATGTGCTTTAGTGGGCTTTGGTTTTCTTTTTTGTGCGAGTGTGTGAGAATGGC TGTGTGGTGAGTGTGAACTTTGTTCTGTGATCATAGAAAGGGTATTTTAGGCTGCAGGGGAG GGCAGGGCTGGGGACCGAAGGGGACAAGTTCCCCTTTCATCCTTTGGTGCTGAGTTTTCTGT AACCCTTGGTTGCCAGAGATAAAGTGAAAAGTGCTTTAGGTGAGATGACTAAATTATGCCTC 

# FIGURE 93

MDLAGLLKSQFLCHLVFCYVFIASGLIINTIQLFTLLLWPINKQLFRKINCRLSYCISSQLV
MLLEWWSGTECTIFTDPRAYLKYGKENAIVVLNHKFEIDFLCGWSLSERFGLLGGSKVLAKK
ELAYVPIIGWMWYFTEMVFCSRKWEQDRKTVATSLQHLRDYPEKYFFLIHCEGTRFTEKKHE
ISMQVARAKGLPRLKHHLLPRTKGFAITVRSLRNVVSAVYDCTLNFRNNENPTLLGVLNGKK
YHADLYVRRIPLEDIPEDDDECSAWLHKLYQEKDAFQEEYYRTGTFPETPMVPPRRPWTLVN
WLFWASLVLYPFFQFLVSMIRSGSSLTLASFILVFFVASVGVRWMIGVTEIDKGSAYGNSDS
KQKLND

## FIGURE 94

GCGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAA GTAAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTA TACAATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAG GCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGT AAACTTGCAGGAGCATTTTTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAA TAACAGAAAGCTGCTCTACTCATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTT TTTCACAGGGTACCTTTAGTGGTTGCCAATCTGGGCATGTCTGAACAACTGGGTTATAAAAC TGTATCAGGTTCCTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAAT TTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTA CAAGAGGAATTAAAGAGTATATGCAAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACT AGTAAAGGATGTAAACAGATTAAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGG CAGCAAGAGAGAAGACATCCAAAAAGACCCTCAGGAGAACATTTTTCTTTGTCAGGCATTA CGGACCTTTTTTCCAAATTCTGAATTTCTTCATTCATGTGTTATGTCTTTAAAAAAATAGACA TGTTTCTAAAAGTAGCTGTAACTACAACCACCATCTCGATGTAGTAGACAATCTGACCTTAA TGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAGCAT AAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTAGATACACAAGA CAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCC CAGAAACAGATGAAGAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACA  $\verb"TTT" \underline{\textbf{TGA}} \texttt{TCCTTTTAACCTTACAAGGAGATTTTTTTTTTTTGGCTGATGGGTAAAGCCAAACAT"$ TTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTCAC CTGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAAC ATCAGATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCT TACACAGACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAA TGGGCAACGTATTGAGACCATGTCTATTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTAT TTTCAAAATATGGAAAGAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCTCCTTAAG TGATACTTTTTTAGAAGTACATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCA 

# FIGURE 95

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQ
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH
FSNQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSC
MSTGFSRAVQTHSSKFFEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN
RLKREIEKRRGAQIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS
CNYNHHLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDDRWQFKRSRLLDTQDKRSKA
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

### FIGURE 96

 $\tt CCAAGCAGCGCGCGCAGCCGCCGCCGCCCACACCCTCTGCGGTCCCCGCGGCGCCTGCCACCCTTCCCT$  $\tt CCTTCCCGGGTCCCCGCCTCGCCGGCCAGTCAGCTTGCCGGGTTCGCTGCCCGCGAAACCCCGAGGTCACCA$ GGGGACCGTTGCCTGACGCGAGGCCCAGCTCTACTTTTCGCCCCGCGTCTCCTCCGCCTGCTCGCCTCTTCCAC CGCCGCTTCCCGTCCGGTCCCAAAGGTGGGAACGCGTCCGCCCCGGCCCGCACC ATG GCACGGTTCGGCTTGCCAAGTGCGACGTCTTTACGTGTCCAAAGGCTTCAACAAGAACGATGCCCCCCTCCACGAGATCAACGGTGATCAT TTGAAGATCTGTCCCCAGGGTTCTACCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGA TGATTTCAAAAGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTTGCTTCACGTTACAAGAAGTTTG ATGAATTCTTCAAAGAACTACTTGAAAATGCAGAGAAATCCCTGAATGATATGTTTGTGAAGACATATGGCCAT  $\verb|TTATACATGCAAAATTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGT|$ GAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATGTTCCGCCTGGTGAACTCCCAGT ACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCT CGCAAATTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTCGCTCAAGGCTTAGCGGTTGCGGG AGATGTCGTGAGCAAGGTCTCCGTGGTAAACCCCACAGCCCAGTGTACCCCATGCCCTGTTGAAGATGATCTACT  ${\tt GCTCCCACTGCCGGGGTCTCGTGACTGTGAAGCCATGTTACAACTACTGCTCAAACATCATGAGAGGCTGTTTG}$ GGGTCCTTTCAACATTGAATCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGG ATAATAGTGTTCAAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCCAAGCCCCTCCCAGCTGGACGAATT AGCTGGCACTAGTTTGGACCGACTGGTTACTGATGTCAAGGAGAAACTGAAACAGGCCAAGAAATTCTGGTCCT AAAGGCAAAAGCAGGTACCTGTTTGCAGTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCA GGTTGACACCAGCAAACCAGACATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGA  ${\tt GGCTGTGAGTATCAGCAGTGCCCTTCAGAGTTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGA}$ GAAAGCCGACAGTGCTGGTGTCCTGGGGGCACAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGG CGCCTTGTTTCTTACAAGCAAACCAGGGTCCCTTCTTGGCACGTAACATGTACGTATTTCTGAAATATTAAATA GCTGTACAGAAGCAGGTTTTATTTATCATGTTATCTTATTAAAAGAAAAAGCCCAAAAAGC

## FIGURE 97

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQ
GSTCCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMF
VKTYGHLYMQNSELFKDLFVELKRYYVVGNVNLEEMLNDFWARLLERMFRLVNSQYHFTDEY
LECVSKYTEQLKPFGDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHAL
LKMIYCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIES
VMDPIDVKISDAIMNMQDNSVQVSQKVFQGCGPPKPLPAGRISRSISESAFSARFRPHHPEE
RPTTAAGTSLDRLVTDVKEKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLF
AVTGNGLANQGNNPEVQVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFFDISDESSGE
GSGSGCEYQQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

# FIGURE 98

# FIGURE 99

MKVLISSLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRR KFMTVSGLPKKQCPCDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

# FIGURE 100

# FIGURE 101

MAVLVLRLTVVLGLLVLFLTCYADDKPDKPDDKPDDSGKDPKPDFPKFLSLLGTEIIENAVE FILRSMSRSTGFMEFDDNEGKHSSK

## FIGURE 102

GGACGCCAGCGCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCT CAGAGCTGGTCTGCCATGGACATCCTGGTCCCACTCCTGCAGCTGCTGGTGCTTCTTAC CCTGCCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCC CCTACCTGATGGCCGTGCTGACTCCCAAGAGCCAACCGCAAGATGGAGAGCAAGAAACGGGAG CTCTTCAGCCAGATAAAGGGGCCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGG CTGCGGAACCGGAGCCAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACC CAAATCCCCACTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATAT GGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGA AGCTGGGCCTTCATGTGGCAGCAAGTTTTCGAGCCCACCTGGAAACACATTGGGGATGGCTG CTGCCTCACCAGAGAGCCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGG AACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTC AAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACA AGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACC**TAG**CAGAATGAGAGAAGACATT CATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATC CCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCC TGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTC CCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGT CTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCT  $\verb|CTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGAT|\\$ **AACCACG** 

# FIGURE 103

MDILVPLLQLLVLLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQI KGLTGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVV APGEDMRQLADGSMDVVVCTLVLCSVQSPRKVLQEVRRVLRPGGVLFFWEHVAEPYGSWAFM WQQVFEPTWKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFP SSKALICSFPSLQLEQATHQPIYLPLRGT

# FIGURE 104

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGAAGTTGCCTCATCGCAGG CAGATGTTGGGGCTTTGTCCGAACAGCTCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAA ACTAATATTTATATGACAGAAGAAAAAG**ATG**TCATTCCGTAAAGTAAACATCATCATCTTGG TCCTGGCTGTTGCTCTTCTTACTGGTTTTGCACCATAACTTCCTCAGCTTGAGCAGTTTG TTAAGGAATGAGGTTACAGATTCAGGAATTGTAGGGCCTCAACCTATAGACTTTGTCCCAAA TGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCCTGTGGTCATCGCTGCATCTG AAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTCAGCACAACACTCGCTCCAAT GTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCCGGTCCTGGCTCAACAG AAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTAC TTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCA AGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAG AAGATTGTGATTCAGCCTCTACTAAAGTTGTCATCCGTGGAGCAGGAAACCAGTACAATTAC ATTGGCTATCTTGACTATAAAAAGGAAAGAATTCGTAAGCTTTCCATGAAAGCCAGCACTTG CTCATTTAATCCTGGAGTTTTTGTTGCAAACCTGACGGAATGGAAACGACAGAATATAACTA ACCAACTGGAAAAATGGATGAAACTCAATGTAGAAGAGGGACTGTATAGCAGAACCCTGGCT GGTAGCATCACAACACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCGATCC TATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTAA AGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATAT ACTGATGTTTGGGAAAAATGGTATATTCCAGACCCAACAGGCAAATTCAACCTAATCCGAAG ATATACCGAGATCTCAAACATAAAG**TGA**AACAGAATTTGAACTGTAAGCAAGCATTTCTCAG GAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTA GCAAGCCATGGAAAAAGATGTGTCAGCTAGGTAAAGATGACAAACTGCCCTGTCTGGCAGTC AGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTTTTCTT ACATTTTTC

# FIGURE 105

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTDSGIVGPQPIDFVPNALRHAVDGR QEEIPVVIAASEDRLGGAIAAINSIQHNTRSNVIFYIVTLNNTADHLRSWLNSDSLKSIRYK IVNFDPKLLEGKVKEDPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNT ALKPGHAAAFSEDCDSASTKVVIRGAGNQYNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA NLTEWKRQNITNQLEKWMKLNVEEGLYSRTLAGSITTPPLLIVFYQQHSTIDPMWNVRHLGS SAGKRYSPQFVKAAKLLHWNGHLKPWGRTASYTDVWEKWYIPDPTGKFNLIRRYTEISNIK

# FIGURE 106

# FIGURE 107

CGACGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCGTAGTGGGGCTGCGCGCCTGCCACG TGGGCTCCGGGGCCTGCGGCGCGGGCGCTGAGCTGGCAGGGCGGGTCGGGGCGCGGGCTGCA TCCGCATCTCCTCCATCGCCTGCAGTAAGGGCGGCCGCGGCGAGCCTTTGAGGGGAACGACT TGTCGGAGCCCTAACCAGGGGTGTCTCTGAGCCTGGTGGGATCCCCGGAGCGTCACATCACT TTCCGATCACTTCAAAGTGGTTAAAAACTAATATTTATATGACAGAAGAAAAAGATGTCATT CCGTAAAGTAAACATCATCTTGGTCCTGGGCTGTTGCTCTTCTTACTGGTTTTGCAC CATAACTTCCTCAGCTTGAGGCAGTTTGTTAAGGAATGAGGTTACAGATTCAGGAATTGTAG GGCCTCAACCTATAGGACTTTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGA GGAGATTCCTGTGGTCATCGCTGCATCTGAAGACAGGCTTGGGGGGGCCCATTGCAGCTATAA ACAGCATTCAGCACAACACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACA GCAGACCATCTCCGGTCCTGGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAATTG TCAATTTTGACCCTAAACTTTTGGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCC ATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGGTTCCCAGCGCAAAGAAGA CCATATACATGGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCA CTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTCAGCCTCTACTAAAGTTGT CATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAA TTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTTGCAAAC CTGACGGAATGGAAACGACAGAATATAACTAACCAACTGGAAAAATGGATGAAACTCAATGT AGAAGAGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACAACACCTCCTCTGCTTATCG TATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGT GCTGGAAAACGATATTCACCTCAGTTTGTAAAGGCTGCCAAGTTACTCCATTGGAATGGACA TTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGGAAAAATGGTATATTCCA GACCCAACAGGCAAATTCAACCTAATCCGAAGATATACCGAGATCTCAAACATAAAGTGAAA CAGAATTTGAACTGTAAGCAAGCATTTCTCAGGAAGTCCTGGAAGATAGCATGCGTGGGAAG TAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAG ATGTCTCCATCTGCCTTACCAAGTGTTTTCTTACTACAATGCTGAATGACTGGAAAGAAGAA CTGATATGGCTAGTTCAGCTAGCTGGTACAGATAATTCAAAACTGCTGTTGGTTTTAATTTT AAAAA

## FIGURE 108

# FIGURE 109

MGAAISQGALIAIVCNGLVGFLLLLLWVILCWACHSRLPTLTLSLNPVPTPALAPVLRRPHH PRSPAMKAATCCSPEGPWPSLEPRT

# FIGURE 110

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCA GTTCCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTA CTCCCTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAA TCATGTCGGGAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCC ATGATGTTTACCTTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTCATTGGTTAT TTTGGGATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACC TCCACAGGCATCACGGCAGTGCTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATT GACAGTTGAGCTTTTCCAAATCACAAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCC AGCCACTGTGGACATTTGCCATCCTCATTTTCTTCTGGGTCCTCTGGGTGGCTGCTGCTG AGCCTGGGAACTGCAGGAGCTGCCCAGGTTATGGAAGGCGGCCAAGTGGAATATAAGCCCCT TTCGGGCATTCGGTACATGTGGTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCA TCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAGTGGTTACTTGTTATTTCAACAGAAGT AGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAGGATTCCGAGAATCATTGTCA TGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGTCCAGGTACCTGTTCCGA TGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATA TACTACAACTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAAATCT TGTCCAAGAACTCAAGTCACTTTACATCTATTAACTGCTTTGGAGACTTCATAATTTTTCTA GGAAAGGTGTTAGTGGTGTTTTCACTGTTTTTTGGAGGACTCATGGCTTTTAACTACAATCG ATAGTTTTTTATCTGTGTTTTGAAACTGTGCTGGATGCACTTTTCCTGTGTTTTGCTGTTGAT CTGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTT CGTAAAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGA ATGAGGAGGGAACAGAACTCCAGGCCATTGTGAGA**TAG**ATACCCATTTAGGTATCTGTACCT GGAAAACATTTCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTT AGTGAATTTTTTTTTAAAAGACCTAATAAACCCTATTCTTCCTCAAAA

# FIGURE 111

MSGRDTILGLCILALALSLAMMFTFRFITTLLVHIFISLVILGLLFVCGVLWWLYYDYTNDL
SIELDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQ
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI
LACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVM
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKIL
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN
EEGTELQAIVR

## FIGURE 112

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGGACCTCTCCCTGTTTCTTCCTT TGTGGTGAAAATTTTTTGAAAAAAAATTGCCTTCTTCAAACAAGGGTGTCATTCTGATATT  $\texttt{T} \underline{\textbf{ATG}} \texttt{AGGACTGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTTCCTTGTTTTGCTGG}$ TGACTGGAGTACATTCAAACAAAGAAACGGCAAAGAAGATTAAAAAGGCCCAAGTTCACTGTG CCTCAGATCAACTGCGATGTCAAAGCCGGAAAGATCATCGATCCTGAGTTCATTGTGAAATG TCCAGCAGGATGCCAAGACCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACT CCAGTGTGTGTGGCGCTGCCGTACACAGTGGTGTGCTTGATAATTCAGGAGGGAAAATACTT GTTCGGAAGGTTGCTGGACAGTCTGGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTT ATCCCTACCACGATGGAGAGAATCCTTTATCGTCTTAGAAAGTAAACCCAAAAAAGGGTGTAA CCTACCCATCAGCTCTTACATACTCATCATCGAAAAGTCCAGCTGCCCAAGCAGGTGAGACC ACAAAAGCCTATCAGAGGCCACCTATTCCAGGGACAACTGCACAGCCGGTCACTCTGATGCA GCTTCTGGCTGTCACTGTAGCTGTGGCCACCCCCACCACCTTGCCAAGGCCATCCCCTTCTGCTGCTTCTACCACCACCACCACCACCACCACCACCACAGGAGCCAGGAGATGGAT CTCTGGTCCACTGCCACACAAGCAGCCAAAACAGGCCAGAGCTGATCCAAAAGGCAAGATCCAAAAGCAGCCAGAGCTGATCCAAAAGCAGCCAGAGCTGTTCAGGAGCTGATCCAAAAGAAGCAGTTTCCAAAAAGAAGAAATTGAGCACACAGTCTTTGGAGCCAGTATCCCTGGGAGATCCAAAC AATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTTCTAATGT AGGTCGGGCCATCTCTTTGTGACCAAGAACTTCTTTTCCAAAGCCAATGGAAACAGAAGCG GGGCTCCCAATGTGGTGGTGATGGTGGATGGCTGGCCCACGGACAAAGTGGAGGAGGCT TCAAGACTTGCGAGAGAGTCAGGAATCAACATTTTCTTCATCACCATTGAAGGTGCTGCTGA TTGTGACCAACCTCACCAAAGAGTTTGAGATTTCCGACACGGACACGCGCATCGGGGCCGTG AGTGATCACCTATGCGATAGGCGTTGCCTGGGCTGCCCAAGAGGAGCTAGAAGTCATTGCCACCCCGCCAGAGACCACCTCCTTCTTTGTGGACGAGTTTGACAACCTCCATCAGTATGTC CCCAGGATCATCCAGAACATTTGTACAGAGTTCAACTCACAGCCTCGGAAC**TGA**ATTCAGAG 

# FIGURE 113

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKC
PAGCQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSL
SLPRWRESFIVLESKPKKGVTYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQ
LLAVTVAVATPTTLPRPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQNRPRADPGIQ
RQDPSGAAFQKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFR
IQKQLLADVAQALDIGPAGPLMGVVQYGDNPATHFNLKTHTNSRDLKTAIEKITQRGGLSNV
GRAISFVTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAE
NEKQYVVEPNFANKAVCRTNGFYSLHVQSWFGLHKTLQPLVKRVCDTDRLACSKTCLNSADI
GFVIDGSSSVGTGNFRTVLQFVTNLTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDI
LNAIKRVGYWSGGTSTGAAINFALEQLFKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKG
VITYAIGVAWAAQEELEVIATHPARDHSFFVDEFDNLHQYVPRIIQNICTEFNSQPRN

## FIGURE 114

CAGGATGAACTGGTTGCAGTGGCTGCTGCTGCGGGGGGCGCTGAGAGGACACGAGCTCT**A** TGCCTTTCCGGCTGCTCATCCCGCTCGGCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGT GCGCCAGGTCCCGACGCCCCGCCCAGATCCCGCCCACTACAGTTTTTCTCTGACTCTAAT TGATGCACTGGACACCTTGCTGATTTTGGGGAATGTCTCAGAATTCCAAAGAGTGGTTGAAG  $\tt CGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGGAAGTAGA$ GGCTGGATGGCCCTGTTCCGGGCCTCTCCTGAGAATGGCTGAGGAGGCGGCCCGAAAACTCC AACCCAGGAGAGCCCCTGTCACCTGTACGGCAGGGATTGGGACCTTCATTGTTGAATTTGC CACCCTGAGCAGCCTCACTGGTGACCCGGTGTTCGAAGATGTGGCCAGAGTGGCTTTGATGC GCCTCTGGGAGAGCCGGTCAGATATCGGGCTGGTCGGCAACCACATTGATGTGCTCACTGGC AAGTGGGTGGCCCAGGACGCAGGCATCGGGGCTGGCGTGGACTCCTACTTTGAGTACTTGGT GAAAGGAGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTCCTAGAGTATAACAAAG CCATCCGGAACTACACCCGCTTCGATGACTGGTACCTGTGGGTTCAGATGTACAAGGGGACT GTGTCCATGCCAGTCTTCCAGTCCTTGGAGGCCTACTGGCCTGGTCTTCAGAGCCTCATTGG AGACATTGACAATGCCATGAGGACCTTCCTCAACTACTACACTGTATGGAAGCAGTTTGGGG GGCTCCCGGAATTCTACAACATTCCTCAGGGATACACAGTGGAGAAGCGAGAGGGCTACCCA CTTCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACGGGGGATCCCACCCT CCTAGAACTCGGAAGAGATGCTGTGGAATCCATTGAAAAAATCAGCAAGGTGGAGTGCGGAT TTGCAACAATCAAAGATCTGCGAGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCCTG GTCCACCTTCGACGCGGTGATCACCCCCTATGGGGAGTGCATCCTGGGGGGCTGGGGGGTACA TCTTCAACACAGAAGCTCACCCCATCGACCTTGCCGCCCTGCACTGCTGCCAGAGGCTGAAG GAAGAGCAGTGGGAGGTGGAGGACTTGATGAGGGAATTCTACTCTCAAAACGGAGCAGGTC GAAATTTCAGAAAAACACTGTTAGTTCGGGGCCCATGGGAACCTCCAGCAAGGCCAGGAACAC TCTTCTCACCAGAAAACCATGACCAGGCAAGGGAGGGAAGCCTGCCAAACAGAAGGTCCCA CTTCTCAGCTGCCCCAGTCAGCCCTTCACCTCCAAGTTGGCATTACTGGGACAGGTTTTCCT AGACTCCTCA**TAA**CCACTGGATAATTTTTTTTTTTTTTTTTTTTTTGAGGCTAAACTATAATA AATTGCTTTTGGCTATCATAAAA

( MH Is)

# FIGURE 115

MPFRLLIPLGLLCALLPQHHGAPGPDGSAPDPAHYSFSLTLIDALDTLLILGNVSEFQRVVE
VLQDSVDFDIDVNASVFETNIRVVGGLLSAHLLSKKAGVEVEAGWPCSGPLLRMAEEAARKL
LPAFQTPTGMPYGTVNLLHGVNPGETPVTCTAGIGTFIVEFATLSSLTGDPVFEDVARVALM
RLWESRSDIGLVGNHIDVLTGKWVAQDAGIGAGVDSYFEYLVKGAILLQDKKLMAMFLEYNK
AIRNYTRFDDWYLWVQMYKGTVSMPVFQSLEAYWPGLQSLIGDIDNAMRTFLNYYTVWKQFG
GLPEFYNIPQGYTVEKREGYPLRPELIESAMYLYRATGDPTLLELGRDAVESIEKISKVECG
FATIKDLRDHKLDNRMESFFLAETVKYLYLLFDPTNFIHNNGSTFDAVITPYGECILGAGGY
IFNTEAHPIDLAALHCCQRLKEEQWEVEDLMREFYSLKRSRSKFQKNTVSSGPWEPPARPGT
LFSPENHDQARERKPAKQKVPLLSCPSQPFTSKLALLGQVFLDSS

# FIGURE 116

AAAGTTACATTTTCTCTGGAACTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACTGAGTCTACCA A**ATG**CAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA AACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC ACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG GAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGA TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT GGTCGTGCCACTGTTCGTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCCAGAAGTTAATCAGCTGCAGAAGG GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT ATGAGGGGACAAGTTGTGTTTCCGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT GGCTTGGAGAGCCCACTTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG TGTTGAGTTCACTTCAAGCCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAAGCGGTAAAAAA AAAAAAAAA

# FIGURE 117

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLLMWSPVIAPGE TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTSAW SILKHPFNRNSTILTRPGMEITKDGFHLVIELEDLGPQFEFLVAYWRREPGAEEHVKMVRSG GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMLILV VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

### Important features:

### Signal peptide:

amino acids 1-29

### Transmembrane domain:

amino acids 230-255

## N-glycosylation sites.

amino acids 40-43 and 134-137

### Tissue factor proteins homology.

amino acids 92-119

## Integrins alpha chain protein homology.

amino acids 232-262

# FIGURE 118

## FIGURE 119

 $\tt CGGACGCGTGGGCCGCCACCTCCGGAACAAGCC{\color{red} ATG} GTGGCGGCGACGGTGGCAGCGGCGTG$ GCTGCTCCTGTGGGCTGCGGCCTGCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAGGCGG TCAACATCCGGGGCAAACTGGTGTCGCTGGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTG AATGTGGCCAGCGAGTGCGGCTTCACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCG AGACCTGGGCCCCCACCACTTTAACGTGCTCGCCTTCCCCTGCAACCAGTTTGGCCAACAGG AGCCTGACAGCAACAAGGAGATTGAGAGCTTTGCCCGCCGCACCTACAGTGTCTCATTCCCC ATGTTTAGCAAGATTGCAGTCACCGGTACTGGTGCCCATCCTGCCTTCAAGTACCTGGCCCA GACTTCTGGGAAGGAGCCCACCTGGAACTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGG TGGTAGGGGCTTGGGACCCAACTGTGTCAGTGGAGGAGGTCAGACCCCAGATCACAGCGCTC  $\tt GTGAGGAAGCTCATCCTACTGAAGCGAGAAGACTTA{\color{red}{\textbf{TAA}}}CCACCGCGTCTCCTCCTCCACCA$ CCTCATCCCGCCCACCTGTGTGGGGCTGACCAATGCAAACTCAAATGGTGCTTCAAAGGGAG AGACCCACTGACTCTCCTTTACTCTTATGCCATTGGTCCCATCATTCTTGTGGGGGAA AAATTCTAGTATTTTGATTATTTGAATCTTACAGCAACAAATAGGAACTCCTGGCCAATGAG AGCTCTTGACCAGTGAATCACCAGCCGATACGAACGTCTTGCCAACAAAAATGTGTGGCAAA TAGAAGTATATCAAGCAATAATCTCCCACCCAAGGCTTCTGTAAACTGGGACCAATGATTAC CTCATAGGGCTGTTGTGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGC CAAATAGGAGGCATTCAATGAACATTTTTTTGCATATAAACCAAAAAATAACTTGTTATCAAT AAAAACTTGCATCCAACATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTT GTTATTTCCTCTGTATTATTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCA AACAATACCTCACGATATAAAATAAAAATGAAAGTATCCTCCTCAAAAA

# FIGURE 120

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVVNVASECGFTDQ HYRALQQLQRDLGPHHFNVLAFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAVTGTG AHPAFKYLAQTSGKEPTWNFWKYLVAPDGKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

## FIGURE 121

CGGACGCGTGGGCGGGCCGGGACGCAGGGCAAAGCGAGCC**ATG**GCTGTCTACGTCGGGATGC TGCGCCTGGGGAGGCTGTGCGCCGGGAGCTCGGGGGTGCTGGGGGCCCGGGCCGCCCTCTCT CGGAGTTGGCAGGAAGCCAGGTTGCAGGGTGTCCGCTTCCTCAGTTCCAGAGAGGTGGATCG CATGGTCTCCACGCCCATCGGAGGCCTCAGCTACGTTCAGGGGTGCACCAAAAAGCATCTTA ACAGCAAGACTGTGGGCCAGTGCCTGGAGACCACAGCACAGAGGGTCCCAGAACGAGAGGCC TTGGTCGTCCTCCATGAAGACGTCAGGTTGACCTTTGCCCAACTCAAGGAGGAGGTGGACAA AGCTGCTTCTGGCCTCCTGAGCATTGGCCTCTGCAAAGGTGACCGGCTGGGCATGTGGGGAC CTAACTCCTATGCATGGGTGCTCATGCAGTTGGCCACCGCCCAGGCGGCATCATTCTGGTG TCTGTGAACCCAGCCTACCAGGCTATGGAACTGGAGTATGTCCTCAAGAAGGTGGGCTGCAA GGCCCTTGTGTTCCCCAAGCAATTCAAGACCCAGCAATACTACAACGTCCTGAAGCAGATCT GTCCAGAAGTGGAGAATGCCCAGCCAGGGGCCTTGAAGAGTCAGAGGCTCCCAGATCTGACC ACAGTCATCTCGGTGGATGCCCCTTTGCCGGGGACCCTGCTCCTGGATGAAGTGGTGGCGGC TGGCAGCACACGGCAGCATCTGGACCAGCTCCAATACAACCAGCAGTTCCTGTCCTGCCATG ACCCCATCAACATCCAGTTCACCTCGGGGACAACAGGCAGCCCCAAGGGGGCCACCCTCTCC CACTACAACATTGTCAACAACTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGAC ACCAGAGCAGTTGCGGATGATCCTGCCCAACCCCCTGTACCATTGCCTGGGTTCCGTGGCAG GCACAATGATGTCTGATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGC AAGAAGGCACTGGAGGCCATCAGCAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGAT GTTCGTGGACATTCTGAACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAG GTGTCATTGCTGGGTCCCCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAAT ATGAAGGACCTGGTGGTTGCTTATGGAACCACAGAGAACAGTCCCGTGACATTCGCGCACTT CCCTGAGGACACTGTGGAGCAGAAGGCAGAAAGCGTGGGCAGAATTATGCCTCACACGGAGG CCCGGATCATGAACATGGAGGCAGGGACGCTGGCAAAGCTGAACACGCCCGGGGAGCTGTGC ATCCGAGGGTACTGCGTCATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGT AGATCGTGGGCCGCTCTAAGGATATGATCATCCGGGGTGGTGAGAACATCTACCCCGCAGAG CTCGAGGACTTCTTTCACACACCCGAAGGTGCAGGAAGTGCAGGTGGTGGGAGTGAAGGA CGATCGGATGGGGAAGAGTTTGTGCCTGCATTCGGCTGAAGGACGGGGAGGAGACCACGG TGGAGGAGATAAAAGCTTTCTGCAAAGGGAAGATCTCTCACTTCAAGATTCCGAAGTACATC GTGTTTGTCACAAACTACCCCCTCACCATTTCAGGAAAGATCCAGAAATTCAAACTTCGAGA  ${\tt GCAGATGGAACGACATCTAAATCTG}$   ${\tt TGA}$   ${\tt ATAAAGCAGGCCTGTCCTGGCCGGTTGGCTT}$ GACTCTCTCTGTCAGAATGCAACCTGGCTTTATGCACCTAGATGTCCCCAGCACCCAGTTC TCCATCCCCACATTCCCCTGTCTGTCTTGTGATTTGGCATAAAGAGCTTCTGTTTTCTTT GAAAAAAAAAAAAAAA

# FIGURE 122

MAVYVGMLRLGRLCAGSSGVLGARAALSRSWQEARLQGVRFLSSREVDRMVSTPIGGLSYVQ
GCTKKHLNSKTVGQCLETTAQRVPEREALVVLHEDVRLTFAQLKEEVDKAASGLLSIGLCKG
DRLGMWGPNSYAWVLMQLATAQAGIILVSVNPAYQAMELEYVLKKVGCKALVFPKQFKTQQY
YNVLKQICPEVENAQPGALKSQRLPDLTTVISVDAPLPGTLLLDEVVAAGSTRQHLDQLQYN
QQFLSCHDPINIQFTSGTTGSPKGATLSHYNIVNNSNILGERLKLHEKTPEQLRMILPNPLY
HCLGSVAGTMMCLMYGATLILASPIFNGKKALEAISRERGTFLYGTPTMFVDILNQPDFSSY
DISTMCGGVIAGSPAPPELIRAIINKINMKDLVVAYGTTENSPVTFAHFPEDTVEQKAESVG
RIMPHTEARIMNMEAGTLAKLNTPGELCIRGYCVMLGYWGEPQKTEEAVDQDKWYWTGDVAT
MNEQGFCKIVGRSKDMIIRGGENIYPAELEDFFHTHPKVQEVQVVGVKDDRMGEEICACIRL
KDGEETTVEEIKAFCKGKISHFKIPKYIVFVTNYPLTISGKIQKFKLREQMERHLNL

### Signal Peptide:

amino acids 1-22

## Transmembrane Domains:

amino acids 140-161, 213-229, 312-334

#### Putative AMP-binding Domain Signature:

amino acids 260-271

## N-myristoylation Sites:

amino acids 19-24, 22-27, 120-125, 203-208, 268-273, 272-277, 314-319, 318-323, 379-384, 380-385, 409-413

## N-glycosylation Site:

amino acids 282-285

# FIGURE 123

## FIGURE 124

 ${\tt GAGCAGGACGGAGCC}$ AGGCTGGCTGCTGCTGCTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGCTGCG TGCAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGC GTGGACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGC AGTGCGGGGTTGCGGTTCGGGACTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGGC TTCTGGCGTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTC ACCTCGCGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCCAACGGCGTGGAGTG CTACAGCTGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTCGTGAGCT GCTACAACGCCAGCGATCATGTCTACAAGGGCTGCTTCGACGCCAACGTCACCTTGACGGCA GCTAATGTGACTGTCCTTGCCTGTCCGGGGCTGTGTCCAGGATGAATTCTGCACTCGGGA TGGAGTAACAGGCCCAGGGTTCACGCTCAGTGGCTCCTGTTGCCAGGGGTCCCGCTGTAACT CTGACCTCCGCAACAAGACCTACTTCTCCCCTCGAATCCCACCCCTTGTCCGGCTGCCCCCT CCAGAGCCCACGACTGTGGCCTCAACCACATCTGTCACCACTTCTACCTCGGCCCCAGTGAG ACCCACATCCACCACAAACCCATGCCAGCGCCAACCAGTCAGACTCCGAGACAGGGAGTAG AACACGAGGCCTCCCGGGATGAGGAGCCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGAC CGCAGCAATTCAGGGCAGTATCCTGCAAAAGGGGGGCCCCAGCAGCCCCATAATAAAGGCTG **GA**GCTTCTCCACCTGGAAATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCT CATCACTTCCTGTTCCCACCACTGGACTGGGCTGGCCCAGCCCCTGTTTTTCCAACATTCCC CAGTATCCCCAGCTTCTGCTGCGCTGGTTTGCGGCTTTTGGGAAATAAAATACCGTTGTATAT ATTCTGCCAGGGGTGTTCTAGCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTC TCCGCTTGTCCTCTTGTGATGTTAGGACAGAGTGAGAGAAGTCAGCTGTCACGGGGAAGGTG GGTGGGTGGGACAATGGCTCCCCACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGG GAATCGGTTCCCCATATGTCTTCCTTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCGGTAC CCAATTCGCCCTATAGTGAGTCGTA

# FIGURE 125

MDPARKAGAQAMIWTAGWLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCT EAVGAVETIHGQFSLAVRGCGSGLPGKNDRGLDLHGLLAFIQLQQCAQDRCNAKLNLTSRAL DPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCFDGNVTLTAANVTV SLPVRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRIPPLVRLPPPEPTT VASTTSVTTSTSAPVRPTSTTKPMPAPTSQTPRQGVEHEASRDEEPRLTGGAAGHQDRSNSG QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

# FIGURE 126

CGGGACTCGGCGGGTCCTCCTGGGAGTCTCGGAGGGGACCGGCTGTGCAGACGCC**ATG**GAGT TGGTGCTGGTCTTCCTCTGCAGCCTGCTGGCCCCCATGGTCCTGGCCAGTGCAGCTGAAAAG GAGAAGGAAATGGACCCTTTTCATTATGATTACCAGACCCTGAGGATTGGGGGACTGGTGTT CGCTGTGGTCCTCTTCTCGGTTGGGATCCTCCTTATCCTAAGTCGCAGGTGCAAGTGCAGTT TCAATCAGAAGCCCCGGGCCCCAGGAGATGAGGAAGCCCAGGTGGAGAACCTCATCACCGCC AATGCAACAGAGCCCCAGAAGCAGAAACTGAAGTGCAGCCATCAGGTGGAAGCCTCTGGAA  $\verb|CCTGAGGCGGCTGCTTGAACCTTTGGATGCAAATGTCGATGCT| \textbf{TAA} \\ \texttt{GAAAACCGGCCACTTC} \\$ AGCAACAGCCCTTTCCCCAGGAGAAGCCAAGAACTTGTGTGTCCCCCACCCTATCCCCTCTA ACACCATTCCTCCACCTGATGATGCAACTAACACTTGCCTCCCCACTGCAGCCTGCGGTCCT CCCAGGCAGGGGCTGAGCCACATGGCCATCTGCTCCTCCCTGCCCCCGTGGCCCTCCATCAC CTTCTGCTCCTAGGAGGCTGCTTGTTGCCCGAGACCAGCCCCCTCCCCTGATTTAGGGATGC GTAGGGTAAGAGCACGGGCAGTGGTCTTCAGTCGTCTTGGGGACCTGGGAAGGTTTGCAGCAC TTTGTCATCATTCTTCATGGACTCCTTTCACTCCTTTAACAAAACCTTGCTTCCTTATCCC ACCTGATCCCAGTCTGAAGGTCTCTTAGCAACTGGAGATACAAAGCAAGGAGCTGGTGAGCC CAGCGTTGACGTCAGGCAGGCTATGCCCTTCCGTGGTTAATTTCTTCCCAGGGGCTTCCACG AGGAGTCCCCATCTGCCCCGCCCCTTCACAGAGCGCCCGGGGATTCCAGGCCCAGGGCTTCT ACTCTGCCCCTGGGGAATGTGTCCCCTGCATATCTTCTCAGCAATAACTCCATGGGCTCTGG GACCCTACCCCTTCCAACCTTCCCTGCTTCTGAGACTTCAATCTACAGCCCAGCTCATCCAG GTTGGGGCCAGCACACCGGGATGGATGGAGGGGAGAGCAGAGGCCTTTGCTTCTCTGCCTACG TCCCCTTAGATGGGCAGCAGAGGCAACTCCCGCATCCTTTGCTCTGCCTGTCGGTGGTCAGA GCGGTGAGCGAGGTGGGTTGGAGACTCAGCAGGCTCCGTGCAGCCCTTGGGAACAGTGAGAG GTTGAAGGTCATAACGAGAGTGGGAACTCAACCCAGATCCCGCCCCTCCTGTCCTCTGTGTT CCCGCGGAAACCAACCAAACCGTGCGCTGTGACCCATTGCTGTTCTCTGTATCGTGATCTAT CCTCAACAACAACAGAAAAAAGGAATAAAATATCCTTTGTTTCCT

# FIGURE 127

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRRCK CSFNQKPRAPGDEEAQVENLITANATEPQKQRTEVQPSGGSLWNLRRLLEPLDANVDA

# FIGURE 128

# FIGURE 129

MKIPVLPAVVLLSLLVLHSAQGATLGGPEEESTIENYASRPEAFNTPFLNIDKLRSAFKADE FLNWHALFESIKRKLPFLNWDAFPKLKGLRSATPDAQ

# FIGURE 130

# FIGURE 131

 ${\tt MGVEIAFASVILTCLSLLAAGVSQVVLLQPVPTQETGPKAMGDLSCGFAGHS}$ 

# FIGURE 132

GGGGAATCTGCAGTAGGTCTGCCGGCG**ATG**GAGTGGTGGGCTAGCTCGCCGCTTCGGCTCTG GCTGCTGTTGTTCCTCCTGCCCTCAGCGCAGGGCCGCCAGAAGGAGTCAGGTTCAAAATGGA AAGTATTTATTGACCAAATTAACAGGTCTTTGGAGAATTACGAACCATGTTCAAGTCAAAAC TGCAGCTGCTACCATGGTGTCATAGAAGAGGATCTAACTCCTTTCCGAGGAGGCATCTCCAG GAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATCACTAAGAACA GACTGTACCGGGAAAATGACTGCATGTTCCCCTCAAGGTGTAGTGGTGTTGAGCACTTTATT TTGGAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGATCAATGTACGAGATTATCCTCA GGTTCCTAAATGGATGGAGCCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACC ATGATATCATGTATCCTGCTTGGACATTTTGGGAAGGGGGACCTGCTGTTTGGCCAATTTAT CCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTGGTAAGGTCAGCAGCACAGTG GCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCAGAACGAG ATCCTCTCATTCTTCTGTCTCGGAAAAACCCAAAACTTGTTGATGCAGAATACACCAAAAAC CAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGT GGATCACTGCAAATACAAGTATCTGTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTA AACACCTCTTCCTGTGTGGCTCACTTGTTTTCCATGTTGGTGATGAGTGGCTAGAATTCTTC TATCCACAGCTGAAGCCATGGGTTCACTATATCCCAGTCAAAACAGATCTCTCCAATGTCCA AGAGCTGTTACAATTTGTAAAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGGGGAA GCCAGTTTATTAGGAACCATTTGCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTG AGTGAATACTCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTAT  ${ t TCCCAAAATGTTGAAAACTGAACTA}$   ${ t TAG}$   ${ t TAGTCATAGGACCATAGTCCTCTTTGTGGCA}$ ACAGATCTCAGATATCCTACGGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATA TCTGCTATCAAGCCAAATACCTGGTTTTCCTTATCATGCTGCACCCAGAGCAACTCTTGAGA AAGATTTAAAATGTGTCTAATACACTGATATGAAGCAGTTCAACTTTTTGGATGAATAAGGA CCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTCATTTTCTTAAGACCAATC ACAGCTTGTGCCTCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCA TGTGATGATGCCCTTTGTCCCATTATTTGGAGCAGAAAATTCGTCATTTGGAAGTAGTACAA CTCATTGCTGGAATTGTGAAATTATTCAAGGCGTGATCTCTGTCACTTTATTTTAATGTAGG AAACCCTATGGGGTTTATGAAAAATACTTGGGGATCATTCTCTGAATGGTCTAAGGAAGCGG TAGCCATGCCATGCAATGATGTAGGAGTTCTCTTTTTGTAAAACCATAAACTCTGTTACTCAG GAGGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATTGCAATTGGATT TCAGGTTCCCTTTTTGTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA

# FIGURE 133

MEWWASSPLRLWLLLFLLPSAQGRQKESGSKWKVFIDQINRSLENYEPCSSQNCSCYHGVIE EDLTPFRGGISRKMMAEVVRRKLGTHYQITKNRLYRENDCMFPSRCSGVEHFILEVIGRLPD MEMVINVRDYPQVPKWMEPAIPVFSFSKTSEYHDIMYPAWTFWEGGPAVWPIYPTGLGRWDL FREDLVRSAAQWPWKKKNSTAYFRGSRTSPERDPLILLSRKNPKLVDAEYTKNQAWKSMKDT LGKPAAKDVHLVDHCKYKYLFNFRGVAASFRFKHLFLCGSLVFHVGDEWLEFFYPQLKPWVH YIPVKTDLSNVQELLQFVKANDDVAQEIAERGSQFIRNHLQMDDITCYWENLLSEYSKFLSY NVTRRKGYDQIIPKMLKTEL

# FIGURE 134

 $\texttt{CACCCTCCATTTCTCGCC} \underline{\textbf{ATG}} \texttt{GCCCCTGCACTGCTCCTGATCCCTGCTCCTCGCCTCTT}$ TCATCCTGGCCTTTGGCACCGGAGTGGAGTTCGTGCGCTTTACCTCCCTTCGGCCACTTCTT CCGCAGCATCCTTGCCCCCCTGGCATGGGATCTGGGGCTCCTGCTTCTATTTGTTGGGCAGC ACAGCCTCATGGCAGCTGAAAGAGTGAAGGCATGGACATCCCGGTACTTTGGGGTCCTTCAG AGGTCACTGTATGTGGCCTGCACTGCCCTGGCCTTGCAGCTGGTGATGCGGTACTGGGAGCC CATACCCAAAGGCCCTGTGTTGTGGGAGGCTCGGGCTGAGCCATGGGCCACCTGGGTGCCGC TCCTCTGCTTTGTGCTCCATGTCATCTCCTGGCTCCTCATCTTTAGCATCCTTCTCGTCTTT GACTATGCTGAGCTCATGGGCCTCAAACAGGTATACTACCATGTGCTGGGGCTGGGCCGAGCC TCTGGCCCTGAAGTCTCCCCGGGCTCTCAGACTCTTCTCCCACCTGCGCCACCCAGTGTGTG TGGAGCTGCTGACAGTGCTGTGGGTGCCTACCCTGGGCACGGACCGTCTCCTTGCT TTCCTCCTTACCCTCTACCTGGGCCTGGCTCACGGGCTTGATCAGCAAGACCTCCGCTACCT  ${\tt CCGGGCCCAGCTACAAAGAAAACTCCACCTGCTCTCTCGGCCCCAGGATGGGGAGGCAGAG{f T}}$ **GA**GGAGCTCACTCTGGTTACAAGCCCTGTTCTTCCTCTCCCACTGAATTCTAAATCCTTAAC ATCCAGGCCCTGGCTGCTTCATGCCAGAGGCCCAAATCCATGGACTGAAGGAGATGCCCCTT CTACTACTTGAGACTTTATTCTCTGGGTCCAGCTCCATACCCTAAATTCTGAGTTTCAGCCA CTGAACTCCAAGGTCCACTTCTCACCAGCAAGGAAGAGTGGGGTATGGAAGTCATCTGTCCC TTCACTGTTTAGAGCATGACACTCTCCCCCTCAACAGCCTCCTGAGAAGGAAAGGATCTGCC CTGACCACTCCCCTGGCACTGTTACTTGCCTCTGCGCCTCAGGGGTCCCCTTCTGCACCGCT GGCTTCCACTCCAAGAAGGTGGACCAGGGTCTGCAAGTTCAACGGTCATAGCTGTCCCTCCA GGCCCCAACCTTGCCTCACCACTCCCGGCCCTAGTCTCTGCACCTCCTTAGGCCCTGCCTCT GGGCTCAGACCCCAACCTAGTCAAGGGGATTCTCCTGCTCTTAACTCGATGACTTGGGGCTC 

# FIGURE 135

MAPALLLIPAALASFILAFGTGVEFVRFTSLRPLLGGIPESGGPDARQGWLAALQDRSILAP LAWDLGLLLLFVGQHSLMAAERVKAWTSRYFGVLQRSLYVACTALALQLVMRYWEPIPKGPV LWEARAEPWATWVPLLCFVLHVISWLLIFSILLVFDYAELMGLKQVYYHVLGLGEPLALKSP RALRLFSHLRHPVCVELLTVLWVVPTLGTDRLLLAFLLTLYLGLAHGLDQQDLRYLRAQLQR KLHLLSRPQDGEAE

## Signal sequence:

amino acids 1-13

### Transmembrane domains:

amino acids 58-76, 99-113, 141-159, 203-222

### N-myristoylation sites:

amino acids 37-43, 42-48, 229-235

# FIGURE 136

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGA AGAAATTGCCAAACCATGTCTTTTTTTCTGTTTTCAGAGTAGTTCACAACAGATCTGAGTGT TTTAATTAAGCATGGAATACAGAAAACAACAAAAAACTTAAGCTTTAATTTCATCTGGAATT TCACGTGGTGCTCTCCGACTACTCACCCCGAGTGTAAAGAACCTTCGGCTCGCGTGCTTCTG AGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCCTTCCGAGTAGGATGTCACTGAGATCC CTCAAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCT TCCCCACTACAATGTGATAGAACGCGTGAACTGGATGTACTTCTATGAGTATGAGCCGATTT ACAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAACTGCTCTCATCAAAATCCATTT TTGGGGTGAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAG AGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGT GACATAATCCGACAAGATTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTATGGC TTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGAGAAGTTT TTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATAT TTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTTGGGTTATATAA TGTCCAGAGATTTGGTGCCAAGGATCTATGAAATGATGGGTCACGTAAAACCCATCAAGTTT GAAGATGTTTATGTCGGGATCTGTTTGAATTTATTAAAAGTGAACATTCATATTCCAGAAGA CAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTTTGGCAGGTCATGCTAAGGAACACC ACATGCCATTAT**TAA**CTTCACATTCTACAAAAAGCCTAGAAGGACAGGATACCTTGTGGAAA GTGTTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTCAGTGTGCTGGCTTACACTG AACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTC AGGCCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAA GAAATTAATAGGACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAAGGG AACAATGTAGTCACTTGAAGGTTTTGTGTATATCTTATGTGGATTACCAATTTAAAAATATA TGTAGTTCTGTGTCAAAAAACTTCTTCACTGAAGTTATACTGAACAAAATTTTACCTGTTTT TGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTCACAGTTATTATTATTTAAAATTA CTTCAACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAAGGATAGTGAAT CATTCTTTACATGCAAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCAC TCCATTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTTAAAT ATTTTACTGTGGTAATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

# FIGURE 137

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQD FHFTLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFFLLGQEAEK EDKMLALSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTDTDVFIN TGNLVKYLLNLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRD LVPRIYEMMGHVKPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHG FSSKEIITFWQVMLRNTTCHY

# FIGURE 138

# FIGURE 139

 $\label{totagligy} {\tt MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNNGWDSWNS\\ {\tt IWDYGNGFAATRLFQKKTCIVHKMNKEVMPSIQSLDALVKEKKLQGKGPGGPPPKGLMYSVN}\\ {\tt PNKVDDLSKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCGDTVEN}$ 

## Signal Peptide:

amino acids 1-20

## N-myristoylation Sites:

amino acids 67-72, 118-123, 163-168

## Flavodoxin protein homology:

amino acids 156-174

# FIGURE 140

CATTTCTGAAACTAATCGTGTCAGAATTGACTTTGAAAAGCATTGCTTTTTACAGAAGTATA TTAACTTTTTAGGAGTAATTTCTAGTTTGGATTGTAATATGAAATAATTTAAAAGGGCTTCG CTCATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTCTTATTGCTTACTGATTTTTT TTGTTCGATTTCAACCAGAGACTATAGCATGTGCTTGCATCTACCTTGCAGCTAGAGCACTT CAGATTCCGTTGCCAACTCGTCCCCATTGGTTTCTTCTTTTTTGGTACTACAGAAGAGGAAAT CCAGGAAATCTGCATAGAAACACTTAGGCTTTATACCAGAAAAAAGCCAAACTATGAATTAC TGGAAAAAGAAGTAGAAAAAAGAAAGTAGCCTTACAAGAAGCCAAATTAAAAGCAAAGGGA TTGAATCCGGATGGAACTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCCAGCCTCCAAGCC ATCATCACCAAGAGAAGTAAAAGCTGAAGAGAAATCACCAATCTCCATTAATGTGAAGACAG TCAAAAAAGAACCTGAGGATAGACAACAGGCTTCCAAAAGCCCTTACAATGGTGTAAGAAAA TTCTAGATCACATACTCCAAGAAGACACTATAATAATAGGCGGAGTCGATCTGGAACATACA GCTCGAGATCAAGAAGCAGGTCCCGCAGTCACAGTGAAAGCCCTCGAAGACATCATAATCAT GGTTCTCCTCACCTTAAGGCCAAGCATACCAGAGATGATTTAAAAAGTTCAAACAGACATGG TCATAAAAGGAAAAATCTCGTTCTCGATCTCAGAGCAAGTCTCGGGATCACTCAGATGCAG CCAAGAAACACAGGCATGAAAGGGGACATCATAGGGACAGGCGTGAACGATCTCGCTCCTTT  ${\tt CTGA} {\tt CTTTCTTTTGAGCCTGCATCAGTTCTTGGTTTTTGCCTATCTACAGTGTGATGT$ ATGGACTCAATCAAAAACATTAAACGCAAACTGATTAGGATTTGATTTCTTGAAACCCTCTA GGTCTCTAGAACACTGAGGACAGTTTCTTTTGAAAAGAACTATGTTAATTTTTTTGCACATT AAAATGCCCTAGCAGTATCTAATTAAAAACCATGGTCAGGTTCAATTGTACTTTATTATAGT TGTGTATTGTTTATTGCTATAAGAACTGGAGCGTGAATTCTGTAAAAATGTATCTTATTTTT ATACAGATAAAATTGCAGACACTGTTCTATTTAAGTGGTTATTTGTTTAAATGATGGTGAAT ACTTTCTTAACACTGGTTTGTCTGCATGTGTAAAGATTTTTACAAGGAAATAAAATACAAAT 

# FIGURE 141

MNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLLFGTTEEEIQEICIETLRLY TRKKPNYELLEKEVEKRKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEEK SPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSRTRSRSRSHTPRRHYN NRRSRSGTYSSRSRSRSRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHKRKKSRSRSQ SKSRDHSDAAKKHRHERGHHRDRRERSRSFERSHKSKHHGGSRSGHGRHRR

# FIGURE 142

# FIGURE 143

 $\mathsf{GGCACGAGGCCTCGTGCCAAGCTTGGCACGAGGGTGCACCGCGTTCTCGCACGCGTC}$ GGTCCTCGGAGTACAGCTGGTGACCCTGCTCACTGCCACCCTCATGCACAGGCTGGCGC CACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAACGGCAGTTTGTTCCGATACAAGCACCCG TCTGAGGAGGAGCTTCGGGCCCTGGCGGGGAAGCCGAGGCCCAGAGGCAGGAAAGAGCGGTG GGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCCGAGATGCCCCGTTCCAGCTGG AGACCTGCCCCTCACGACCGTGGATGCCCTGGTCCTGCGCTTCTTCCTGGAGTACCAGTGG TTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTTCACAGAGGCCTACTACAT GCTGGGACCAGCCAAGGAGACTAACATTGCTGTTCTGGTGCCTGCTCACGGTGACCTTCT CCATCAAGATGTTCCTGACAGTGACACGGCTGTACTTCAGCGCCGAGGAGGGGGGTGAGCGC TCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGGCCATGCTGGTGCAAGTGGTGCG GGAGGAGACCCTCGAGCTGGGCCTGGAGCCTGGTCTGGCCAGCATGACCCAGAACTTAGAGC CACTTCTGAAGAAGCAGGGCTGGGACTGGGCGCTTCCTGTGGCCAAGCTGGCTATCCGCGTG GGACTGGCAGTGGTGGGCTCTGTGCTGGGTGCCTTCCTCACCTTCCCAGGCCTGCGGCTGGC CCAGACCCACCGGGACGCACTGACCATGTCGGAGGACAGACCCATGCTGCAGTTCCTCCTGC ACACCAGCTTCCTGTCTCCCCTGTTCATCCTGTGGCTCTGGACAAAGCCCCATTGCACGGGAC TTCCTGCACCAGCCGCCGTTTGGGGAGACGCGTTTCTCCCTGCTGTCCGATTCTGCCTTCGA ACCTGCAGGCCTACCTGTGCCTGGCCAAGGCCCGGGTGGAGCAGCTGCGAAGGGAGGCTGGC CGCATCGAAGCCCGTGAAATCCAGCAGAGGGTGGTCCGAGTCTACTGCTATGTGACCGTGGT GAGCTTGCAGTACCTGACGCCGCTCATCCTCACCCTCAACTGCACACTTCTGCTCAAGACGC TGGGAGGCTATTCCTGGGGCCTGGGCCCAGCTCCTCTACTATCCCCCGACCCATCCTCAGCC AGCGCTGCCCCCATCGGCTCTGGGGAGGACGAAGTCCAGCAGACTGCAGCGCGGATTGCCGG GGCCCTGGGTGGCCTGCTTACTCCCCTCTTCCTCCGTGGCGTCCTGGCCTACCTCATCTGGT GGACGGCTGCCTGCCAGCTCCCAGCCTTTTCGGCCTCTACTTCCACCAGCACTTGGCA  $\mathsf{GGCTCC}$ CTAGCCTGCCCCTCTGTTTGCGCCCCCGTGTCCCCAGCTGCAAGGTGGGGCCGGACTCCCC GGCGTTCCCTTCACCACAGTGCCTGACCCGCGGCCCCCTTGGACGCCGAGTTTCTGCCTCA GAACTGTCTCCTGGGCCCAGCAGCATGAGGGTCCCGAGGCCATTGTCTCCGAAGCGTATG TGCCAGGTTTGAGTGGCGAGGGTGATGCTGGCTGCTCTTCTGAACAAATAAAGGAGCATGCC **GATTTTTAA** 

# FIGURE 144

MAVLGVQLVVTLLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEEELRALAGKPRPRGRKE RWANGLSEEKPLSVPRDAPFQLETCPLTTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY YMLGPAKETNIAVFWCLLTVTFSIKMFLTVTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV VREETLELGLEPGLASMTQNLEPLLKKQGWDWALPVAKLAIRVGLAVVGSVLGAFLTFPGLR LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPPFGETRFSLLSDSA FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT VVSLQYLTPLILTLNCTLLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQQTAARI AGALGGLLTPLFLRGVLAYLIWWTAACQLLASLFGLYFHQHLAGS

## FIGURE 145

## FIGURE 146

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTTATTAACGTGGCTT AATCTGAAGGTTCTCAGTCAAATTCTTTGTGATCTACTGATTGTGGGGGGCCATGGCAAGGTTTGCTTAAAGGAGC  $\tt TTGGCTGGTTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGA\\ {\color{red} \underline{ATG}} {\color{blue} AAGG}$ CGCTTCTGTTGCTGGTCTTGCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTG  ${\tt TCCAGACGGCTGTGCGAGCCTCACAGCCACGGCTCCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAA}$ GTGGACTCTGGCCGGAGCAACCGAACTAGGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAA AAAAATAAATCGAGCTTTGAGTGTTCTTCGAAGGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGG GCAGGGAAAATTCTGAAAACACCACTGCCCCTGAAGTCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAA ATTACCAGCATCAAGATCAATCGAGTAGATCCCAGTGAAAGCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAAC GAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACTACGCTGTGCGTCTCCTGCGG CAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCCGCAGCAGGAACAATGGACAGGCCCC GGATGCCTACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGGAA  $\tt CTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACACT$  $\verb|CCCAAGCCCTCCATCCTACAATTACTTGTCATGAGAAGGTGGTAAATATCCAAAAAGACCCCGGTGAATCTCT| \\$ GAGGTCAGCCGGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGT GTGACTGGTCCCCATCCTGGGTCATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTA TTTCATCAAATCCATTGTTGAAGGAACACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTG ATTACTCTAACTATTGTTTCTTGGCCTGGCACTTTTTTATAGAATCAATGATGGGTCAGAGGAAAACAGAAAAA AAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAAATTA AAACTACTAGTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTACAACATTGTTTATATTTTTTCTATTCAAT AAAAAGCCCTAAAACAACTAAAATGATTGATTTGTATACCCCACTGAATTCAAGCTGATTTAAATTTAAAATTT GGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTTAATTTACAGCTAAAATATTTTTAAAAATGCA TTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAATATTTTTCAGAAGTTAAA

## FIGURE 147

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASLTAT

APSPEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRS

FKKINRALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDP

SESLSIRLVGGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLL

RQPCQVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGV

FIFNVLDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAAHLIQASERRVHLVVSRQVRQRS

PDIFQEAGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHRE

WDLPIYVISVEPGGVISRDGRIKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEV

KEYEPQEDCSSPAALDSNHNMAPPSDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIV

GGYEEYNGNKPFFIKSIVEGTPAYNDGRIRCGDILLAVNGRSTSGMIHACLARLLKELKGRI

TLTIVSWPGTFL

## FIGURE 148

# FIGURE 149

MKILVAFLVVLTIFGIQSHGYEVFNIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT IFDYKHGYIASRVLSRRACFILKMDHQNIPPLNNLQWYIYEKQALDNMFSNKYTWVKYNPLE SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

## FIGURE 150

ATGGGGCTCCCTGGGCTGTTCTGCTTGGCCGTGCTGCCAGCAGCTTCTCCAAGGCACG GGAGGAAGAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTTCCCCAAAG GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCATCACCTATTCCCTC TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG CCAGTGTCTGAGCTGCGGGCCAACTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGGA GATGATCTGCCAGGCGTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAAGGATG AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA GCCCCATCCTTGCCTTGCCGCTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGGAGTTTGGG  $\mathsf{GGGTTCAGGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCAGCCATG\mathbf{TAG}$   $\mathsf{AATGAACC}$ GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCACTGTTCGTATTTGGA 

# FIGURE 151

MGLPGLFCLAVLAASSFSKAREEEITPVVSIAYKVLEVFPKGRWVLITCCAPQPPPPITYSL CGTKNIKVAKKVVKTHEPASFNLNVTLKSSPDLLTYFCRASSTSGAHVDSARLQMHWELWSK PVSELRANFTLQDRGAGPRVEMICQASSGSPPITNSLIGKDGQVHLQQRPCHRQPANFSFLP SQTSDWFWCQAANNANVQHSALTVVPPGGDQKMEDWQGPLESPILALPLYRSTRRLSEEEFG GFRIGNGEVRGRKAAAM

#### Signal Peptide:

amino acids 1-18

#### N-glycosylation Sites:

amino acids 86-89, 132-135, 181-184

## FIGURE 152

GGTCCTTA**ATG**GCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTG CTGTCCGGCTGGTCCCGGGCTGGGCCGACCCTCACTCTCTTTGCTATGACATCACCGT CATCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGA CTTTTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAAA CTAAATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACT TACAGAGCAACTGCGTGACATTCAGCTGGAGAATTACACACCCAAGGAACCCCTCACCCTGC AGGCAAGGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGT TTCGATGGGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCC TGGAGCCAGAAAGATGAAAGAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATT ACTTCTCAATGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACACCC CTGGAGCCAAGTGCAGGAGCACCACTCGCCATGTCCTCAGGCACAACCCAACTCAGGGCCAC AGCCACCACCTCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCA TCTGAGGAGAGTCCTTTAGAGTGACAGGTTAAAGCTGATACCAAAAGGCTCCTGTGAGCACG GTCTTGATCAAACTCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGT GGCCTCCAGCAGATCATGATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCCTT TTGCCAACAATTTTACCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTTGGTGCTACC GTCAGTAAAATAATCACGTTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAG AATTTTTAAATTATTTAATAAGAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTAT 

## FIGURE 153

MAAAAATKILLCLPLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTFL HYDCGNKTVTPVSPLGKKLNVTTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQAR MSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFS MGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

#### Important features:

#### Signal peptide:

amino acids 1-25

#### Transmembrane domain:

amino acids 224-246

#### N-glycosylation site.

amino acids 68-72, 82-86

#### N-myristoylation site.

amino acids 200-206, 210-216

#### Amidation site.

amino acids 77-81

## FIGURE 154

# FIGURE 155

 ${\tt MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHELWLLIKREFGFYSKSQYRTWQKKLA}\\ {\tt EDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGGQPTEQHFWARL}$ 

## FIGURE 156

 $\tt GTTCTCCTTTCCGAGCCAAAATCCCAGGCGATGGTGAATTATGAACGTGCCACACC \\ \underline{\textbf{ATG}} \textbf{AAG} \textbf{AAG}$ CTCTTGTGGCAGGTAACTGTGCACCACCACCTGGAATGCCATCCTGCTCCCGTTCGTCTA CCTCACGGCGCAAGTGTGGATTCTGTGTGCAGCCATCGCTGCTGCCGCCTCAGCCGGGCCCC AGAACTGCCCCTCCGTTTGCTCGTGCAGTAACCAGTTCAGCAAGGTGGTGTGCACGCGCCGG GGCCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACCTCATGGAGAA CAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCCTGCAGT TGGGCAGGAACTCCATCCGGCAGATTGAGGTGGGGGCCTTCAACGGCCTGGCCAGCCTCAAC ACCCTGGAGCTGTTCGACAACTGGCTGACAGTCATCCCTAGCGGGGCCTTTGAATACCTGTC CAAGCTGCGGGAGCTCTGGCTTCGCAACAACCCCATCGAAAGCATCCCCTCTTACGCCTTCA ACCGGGTGCCCTCCCTCATGCGCCTGGACTTGGGGGAGCTCAAGAAGCTGGAGTATATCTCT GAGGGAGCTTTTGAGGGGCTGTTCAACCTCAAGTATCTGAACTTGGGCATGTGCAACATTAA AGACATGCCCAATCTCACCCCCCTGGTGGGGGCTGGAGGAGCTGGAGATGTCAGGGAACCACT TCCCTGAGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTCATG CAACTTGGCCCACAATAACCTCTCTTCTTTGCCCCATGACCTCTTTACCCCGCTGAGGTACC TGGTGGAGTTGCATCTACACCACAACCCTTGGAACTGTGATTGTGACATTCTGTGGCTAGCC TGGTGGCTTCGAGAGTATATACCCACCAATTCCACCTGCTGTGGCCGCTGTCATGCTCCCAT GCACATGCGAGGCCGCTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCT TCATCATGGACGCACCTCGAGACCTCAACATTTCTGAGGGTCGGATGGCAGAACTTAAGTGT CGGACTCCCCTATGTCCTCCGTGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGC CTCCCGCCACCAAGGATCTCTGTCCTCAACGACGCCACCTTGAACTTTTCCCACGTGCTGC TTTCAGACACTGGGGTGTACACATGCATGGTGACCAATGTTGCAGGCAACTCCAACGCCTCG GCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAACTACAGCTTCTTCACCACAGT AACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACGCGAAAGTACAAGCCTGTTCCTA CCACGTCCACTGGTTACCAGCCGGCATATACCACCTCTACCACGGTGCTCATTCAGACTACC CGTGTGCCCAAGCAGGTGGCAGTACCCGCGACAGACACCACTGACAAGATGCAGACCAGCCT GGATGAAGTCATGAAGACCACCAAGATCATCATTGGCTGCTTTGTGGCAGTGACTCTGCTAG CTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGCGGAGTACAGTC ACAGCCGCCCGGACTGTTGAGATAATCCAGGTGGACGAAGACATCCCAGCAGCAACATCCGC AGCAGCAACAGCAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCCACAATTC ATGACCATATTAACTACAACACCTACAAACCAGCACATGGGGCCCACTGGACAGAAAACAGC CTGGGGAACTCTCTGCACCCACAGTCACCACTATCTCTGAACCTTATATAATTCAGACCCA TACCAAGGACAAGGTACAGGAAACTCAAATA**TGA**CTCCCCTCCCCCAAAAAAACTTATAAAAT GCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTTTTCTTGTA TATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAAACAGATTATATTAAAATTTAAAGA CAAAAAGTCAAAACA

## FIGURE 157

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCT
RRGLSEVPQGIPSNTRYLNLMENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLAS
LNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEY
ISEGAFEGLFNLKYLNLGMCNIKDMPNLTPLVGLEELEMSGNHFPEIRPGSFHGLSSLKKLW
VMNSQVSLIERNAFDGLASLVELNLAHNNLSSLPHDLFTPLRYLVELHLHHNPWNCDCDILW
LAWWLREYIPTNSTCCGRCHAPMHMRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAEL
KCRTPPMSSVKWLLPNGTVLSHASRHPRISVLNDGTLNFSHVLLSDTGVYTCMVTNVAGNSN
ASAYLNVSTAELNTSNYSFFTTVTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQ
TTRVPKQVAVPATDTTDKMQTSLDEVMKTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQQRS
TVTAARTVEIIQVDEDIPAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHWTE
NSLGNSLHPTVTTISEPYIIQTHTKDKVQETQI

### FIGURE 158

TGTCGGGAGTGCTGTGAATATGATCAGATTGAGTGCGTCTGCCCCGGAAAGAGGGGAAGTCGTGGGTTATACCAT GCAAGAGCTGCCGAAATGGCTCATGGGGGGGTACCTTGGATGACTTCTATGTGAAGGGGTTCTACTGTGCAGAG TGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTGGCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTT GTTGGAAAGCTATCCCCTAAATGCTCACTGTGAATGGACCATTCATGCTAAACCTGGGTTTGTCATCCAACTAA GATTTGTCATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAGGTTCGTGATGGAGACAAC CGCGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAGGATCCTCACT CCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGCAT GCTCCTCATCCCCTTGTTTCCATGACGGCACGTGCGTCCTTGACAAGGCTGGATCTTACAAGTGTGCCTGCTTG TGGGTACCAGAAAATAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAATTGGCACCGTGGTGTCTT TCTTTTGTAACAACTCCTATGTTCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAATGGAGAGTGGTCAGGG AAACAGCCCATCTGCATAAAAGCCTGCCGAGAACCAAAGATTTCAGACCTGGTGAGAAGGAGAGTTCTTCCGAT CCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGGATACCAACATCTGCATACCCAGCTCCAG TATGAGTGCATCTCACCCTTCTACCGCCGCCTGGGCAGCAGCAGGAGGACATGTCTGAGGACTGGGAAGTGGAG TGGGCGGCACCATCCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGACCCAAGGGTTGC GCTGGCCGTGGCAGCCATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGCGTGG GAAGGTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTTTGGGGAAATTCTACCGGGATGATGACCGGGATG AGAAGACCATCCAGAGCCTACAGATTTCTGCTATCATTCTGCATCCCAACTATGACCCCATCCTGCTTGATGCT GCCCTGGCTTCAAGAACGACACTGCGCTCTGGGGTGGTCAGTGTGGTGGACTCGCTGTGTGAGGAGCAG CATGAGGACCATGGCATCCCAGTGAGTGTCACTGATAACATGTTCTGTGCCAGCTGGGAACCCACTGCCCCTTC TGATATCTGCACTGCAGAGACAGGAGGCATCGCGGCTGTGTCCTTCCCGGGACGAGCATCTCCTGAGCCACGCT GGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCACTGCCTTCACCAAG  ${ t GTGCTGCCTTTTAAAGACTGGATTGAAAGAAATATGAAA}{{ t TGA}}$ CTGTGCCAGGGCTTCTGACTTCAGGGACAAAACTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGAT ATATACAAAACCTCTCCACTCCACTGACCTGGTGGTCTTCCCCAACTTTCAGTTATACGAATGCCATCAGCTTG ACCAGGGAAGATCTGGGCTTCATGAGGCCCCTTTTGAGGCTCTCAAGTTCTAGAGAGCTGCCTGTGGGACAGCC 

## FIGURE 159

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAEWNIMCRECCEYDQIECVCPGKREVV
GYTIPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGD
CMRCGQVLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRD
GDNRDGQIIKRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDG
TCVLDKAGSYKCACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTV
VSFFCNNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMQVQSRETPLH
QLYSAAFSKQKLQSAPTKKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGK
WSGRAPSCIPICGKIENITAPKTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNE
RTVVVAAHCVTDLGKVTMIKTADLKVVLGKFYRDDDRDEKTIQSLQISAIILHPNYDPILLD
ADIAILKLLDKARISTRVQPICLAASRDLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSG
VVSVVDSLLCEEQHEDHGIPVSVTDNMFCASWEPTAPSDICTAETGGIAAVSFPGRASPEPR
WHLMGLVSWSYDKTCSHRLSTAFTKVLPFKDWIERNMK

## FIGURE 160

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGA AGCTTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTCACGTAATAAAAAAC**ATG**GGC TTCAACCTGACTTTCCACCTTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCT GACAGTGGTTGGGTGGGCCACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCCTAAAG CAAAGGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAAACTCTGACT AATGAAGCATCCACGAAGAAGGTAGAACTTGACAACTGTCCTTCTGTGTCTCCTTACCTCAG AGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATC CCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATC CTCGTTCCCCACCGGAACAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTT CCTGCAGAGGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGT TTAATCGAGCCAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGGAC TGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGA GGAGCATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTG GATATTTTGGGGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCT AACAACTACTGGGGATGGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAG AATGAAAATTTCCCGGCCCCTGCCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAG ACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTCACGAGTCTGG AGAACAGATGGGTTGAGTAGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATA AAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTA AGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCCTTTTTGTATTTTCTTAGCAGAGCT CCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGATCATG AGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGAT TATGGGATAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCT CGTCCAAGGTAGAAAGGTACGAAGATACAATACTGTTATTCATTTATCCTGTACAATCATCT GTGAAGTGGTGGTGAGAGGGGGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCA GGACACAGTGAACTTGGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTCGGCTGCAAAGGCAG CAGTAGCTGAGCTGGTTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCCAGGTATGCCT TCCAGTGATGCCCACCAGAGAATACATTCTCTATTAGTTTTTAAAGAGTTTTTGTAAAATGA TGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAGCAAAA

## FIGURE 161

MGFNLTFHLSYKFRLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKT
LTNEASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRV
AILVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEGKKFNRAKLLNVGYLEALKEEN
WDCFIFHDVDLVPENDFNLYKÇEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNG
FSNNYWGWGGEDDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSR
VWRTDGLSSCSYKLVSVEHNPLYINITVDFWFGA

#### Important features:

#### Signal peptide:

amino acids 1-27

#### N-glycosylation sites:

amino acids 4-7, 220-223 and 335-338

#### Xylose isomerase proteins:

amino acids 191-201

### FIGURE 162

CGTGGGCCGGGGTCGCGCAGCGGGCTGTGGGCGCCCCGGAGGAGCGACCGCCGCAGTTCTC GAGCTCCAGCTGCATTCCCTCCGCGTCCGCCCCACGCTTCTCCCGCTCCGGGCCCCGCA**ATG**  $\tt CCCGGCAGGGTGGCCGCAGGCCTGTATGAACTCAATCTCACCACCGATAGCCCTGCCACCA$ CGGGAGCGGTGGTGACCATCTCGGCCAGCCTGGTGGCCAAGGACAACGGCAGCCTGGCCCTG CCCGCTGACGCCCACCTCTACCGCTTCCACTGGATCCACACCCCGCTGGTGCTTACTGGCAA GATGGAGAAGGGTCTCAGCTCCACCATCCGTGTGGTCGGCCACGTGCCCGGGGAATTCCCGG TCTCTGTCTGGGTCACTGCCGCTGACTGCTGGATGTGCCAGCCTGTGGCCAGGGGCTTTGTG GTCCTCCCCATCACAGAGTTCCTCGTGGGGGACCTTGTTGTCACCCAGAACACTTCCCTACC GCAACTTCCTCAAGACCGCCTTGTTTCTCTACAGCTGGGACTTCGGGGACCGGACCCAGATG GTGACTGAAGACTCCGTGGTCTATTATAACTATTCCATCATCGGGACCTTCACCGTGAAGCT CAAAGTGGTGGCGGAGTGGGAAGAGGTGGAGCCGGATGCCACGAGGGCTGTGAAGCAGAAGA CCGGGGACTTCTCCGCCTCGCTGAAGCTGCAGGAAACCCTTCGAGGCATCCAAGTGTTGGGG CCCACCCTAATTCAGACCTTCCAAAAGATGACCGTGACCTTGAACTTCCTGGGGAGCCCTCC TCTGACTGTGTGCTGGCGTCTCAAGCCTGAGTGCCTCCCGCTGGAGGAAGGGGAGTGCCACC CTGTGTCCGTGGCCAGCACAGCGTACAACCTGACCCACACCTTCAGGGACCCTGGGGACTAC TGCTTCAGCATCCGGGCCGAGAATATCATCAGCAAGACACATCAGTACCACAAGATCCAGGT GTGGCCCTCCAGAATCCAGCCGGCTGTCTTTGCTTTCCCATGTGCTACACTTATCACTGTGA TGTTGGCCTTCATCATGTACATGACCCTGCGGAATGCCACTCAGCAAAAGGACATGGTGGAG AACCCGGAGCCACCCTCTGGGGTCAGGTGCTGCTGCCAGATGTGCTGTGGGCCTTTCTTGCT GGAGACTCCATCTGAGTACCTGGAAATTGTTCGTGAGAACCACGGGCTGCTCCCGCCCCTCT ATAAGTCTGTCAAAACTTACACCGTG**TGA**GCACTCCCCCTCCCCACCCCATCTCAGTGTTAA CTGACTGCTGACTTGGAGTTTCCAGCAGGGTGGTGTGCACCACTGACCAGGAGGGGTTCATT TGCGTGGGGCTGTTGGCCTGGATCATCCATCCATCTGTACAGTTCAGCCACTGCCACAAGCC CCTCCCTCTCTGTCACCCCTGACCCCAGCCATTCACCCATCTGTACAGTCCAGCCACTGACA TAAGCCCCACTCGGTTACCACCCCCTTGACCCCCTACCTTTGAAGAGGCTTCGTGCAGGACT TTGATGCTTGGGGTGTTCCGTGTTGACTCCTAGGTGGGCCTGCCCACTGCCCATTCCT CTCATATTGGCACATCTGCTGTCCATTGGGGGGTTCTCAGTTTCCTCCCCCAGACAGCCCTAC CTGTGCCAGAGAGCTAGAAAGAAGGTCATAAAGGGTTAAAAAATCCATAACTAAAGGTTGTAC CACACACACAGAAATATAAACACATGCGTCACATGGGCATTTCAGATGATCAGCTCTGTA TCTGGTTAAGTCGGTTGCTGGGATGCACCCTGCACTAGAGCTGAAAGGAAATTTGACCTCCA AGCAGCCCTGACAGGTTCTGGGCCCGGGCCCTCCCTTTGTGCTTTGTCTCTGCAGTTCTTGC GCCCTTTATAAGGCCATCCTAGTCCCTGCTGGCTGGCAGGGGCCTGGATGGGGGGCAGGACT AATACTGAGTGATTGCAGAGTGCTTTATAAATATCACCTTATTTTATCGAAACCCATCTGTG TGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGATCAGGA AGCCGGGCGTGGTGGGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATG GTGCGAACCCGGGAGCGGAGCTTGCAGTGAGCCCAGATGGCGCCACTGCACTCCAGCCTGA GTGACAGAGCGAGACTCTGTCTCCA

## FIGURE 163

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLA
LPADAHLYRFHWIHTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGF
VVLPITEFLVGDLVVTQNTSLPWPSSYLTKTVLKVSFLLHDPSNFLKTALFLYSWDFGDGTQ
MVTEDSVVYYNYSIIGTFTVKLKVVAEWEEVEPDATRAVKQKTGDFSASLKLQETLRGIQVL
GPTLIQTFQKMTVTLNFLGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYNLTHTFRDPGD
YCFSIRAENIISKTHQYHKIQVWPSRIQPAVFAFPCATLITVMLAFIMYMTLRNATQQKDMV
ENPEPPSGVRCCCQMCCGPFLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

#### Important features of the protein:

#### Signal peptide:

amino acids 1-24

#### Transmembrane domain:

amino acids 339-362

#### N-glycosylation sites.

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

### FIGURE 164

# FIGURE 165

MALSSQIWAACLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTH FPICIFCCGCCHRSKCGMCCKT

### FIGURE 166

CTGTCAGGAAGGACCATCTGAAGGCTGCAATTTGTTCTTAGGGAGGCAGGTGCTGGCCTGGC  $\tt CTGGATCTTCCACC {\color{red} \underline{ATG}} {\color{blue} TTCCTGTTGCTGCCTTTTGATAGCCTGATTGTCAACCTTCTGGGC$ AGTCTCCTTTGGTATCCGCAAACTCTACATGAAAAGTCTGTTAAAAATCTTTGCGTGGGCTA CCTTGAGAATGGAGCGAGGAGCCAAGGAGAAGAACCACCAGCTTTACAAGCCCTACACCAAC GGAATCATTGCAAAGGATCCCACTTCACTAGAAGAAGAGATCAAAGAGATTCGTCGAAGTGG TAGTAGTAAGGCTCTGGACAACACTCCAGAGTTCGAGCTCTCTGACATTTTCTACTTTTGCC GGAAAGGAATGGAGACCATTATGGATGATGAGGTGACAAAGAGATTCTCAGCAGAAGAACTG GAGTCCTGGAACCTGCTGAGCAGAACCAATTATAACTTCCAGTACATCAGCCTTCGGCTCAC GGTCCTGTGGGGGTTAGGAGTGCTGATTCGGTACTGCTTTCTGCTGCCGCTCAGGATAGCAC TGGCTTTCACAGGGATTAGCCTTCTGGTGGTGGGCACAACTGTGGTGGGATACTTGCCAAAT AGCGCTGACAGCCATCATCACCTACCATGACAGGGAAAACAGACCAAGAAATGGTGGCATCT GTGTGGCCAATCATACCTCACCGATCGATGTGATCATCTTGGCCAGCGATGGCTATTATGCC ATGGTGGGTCAAGTGCACGGGGGACTCATGGGTGTGATTCAGAGAGCCATGGTGAAGGCCTG CCCACACGTCTGGTTTGAGCGCTCGGAAGTGAAGGATCGCCACCTGGTGGCTAAGAGACTGA CTGAACATGTGCAAGATAAAAGCAAGCTGCCTATCCTCATCTTCCCAGAAGGAACCTGCATC AATAATACATCGGTGATGATGTTCAAAAAGGGAAGTTTTGAAATTGGAGCCACAGTTTACCC TGTTGCTATCAAGTATGACCCTCAATTTGGCGATGCCTTCTGGAACAGCAGCAAATACGGGA TGGTGACGTACCTGCGGAATGATGACCAGCTGGGCCATTGTCTGCAGCGTGTGGTACCTG CCTCCCATGACTAGAGAGGCAGATGAAGATGCTGTCCAGTTTGCGAATAGGGTGAAATCTGC CATTGCCAGGCAGGAGGACTTGTGGACCTGCTGTGGGATGGGGGCCTGAAGAGGGAGAAGG TGAAGGACACGTTCAAGGAGGAGCAGCAGAAGCTGTACAGCAAGATGATCGTGGGGAACCAC AAGGACAGGAGCCGCTCCTGAGCTGCCTCCAGCTGGGGGCCACCGTGCGGGGTGCCAA  $\tt CGGGCTCAGAGCTGGAGTTGCCGCCGCCGCCCCCACTGCTGTGTCCTTTCCAGACTCCAGGG$  $\tt CTCCCCGGGCTGCTCTGGATCCCAGGACTCCGGCTTTCGCCGAGCCGCAGCGGGATCCCTGT$ GCACCCGGCGCAGCCTACCCTTGGTGGTCTAAACGGATGCTGCTGGGTGTTGCGACCCAGGA CGAGATGCCTTGTTTCTTTTACAATAAGTCGTTGGAGGAATGCCATTAAAGTGAACTCCCCA CCTTTGCACGCTGTGCGGGCTGAGTGGTTGGGGAGATGTGGCCATGGTCTTGTGCTAGAGAT GGCGGTACAAGAGTCTGTTATGCAAGCCCGTGTGCCAGGGATGTGCTGGGGGGCGGCCACCCG CTCTCCAGGAAAGGCACAGCTGAGGCACTGTGGCTGGCTTCGGCCTCAACATCGCCCCCAGC CTTGGAGCTCTGCAGACATGATAGGAAGGAAACTGTCATCTGCAGGGGCTTTCAGCAAAATG GGCCGCTGACTGGGCCATGGGGAGAACGTGTTTCGTACTCCAGGCTAACCCTGAACTCCCC ATGTGATGCGCGCTTTGTTGAATGTGTGTCTCGGTTTCCCCATCTGTAATATGAGTCGGGGG AGGACACATCACGTTCAGTGTTTCAAGTACAGGCCCACAAAACGGGGCACGGCAGGCCTGAG TGA

# FIGURE 167

MFLLLPFDSLIVNLLGISLTVLFTLLLVFIIVPAIFGVSFGIRKLYMKSLLKIFAWATLRME
RGAKEKNHQLYKPYTNGIIAKDPTSLEEEIKEIRRSGSSKALDNTPEFELSDIFYFCRKGME
TIMDDEVTKRFSAEELESWNLLSRTNYNFQYISLRLTVLWGLGVLIRYCFLLPLRIALAFTG
ISLLVVGTTVVGYLPNGRFKEFMSKHVHLMCYRICVRALTAIITYHDRENRPRNGGICVANH
TSPIDVIILASDGYYAMVGQVHGGLMGVIQRAMVKACPHVWFERSEVKDRHLVAKRLTEHVQ
DKSKLPILIFPEGTCINNTSVMMFKKGSFEIGATVYPVAIKYDPQFGDAFWNSSKYGMVTYL
LRMMTSWAIVCSVWYLPPMTREADEDAVQFANRVKSAIARQGGLVDLLWDGGLKREKVKDTF
KEEQQKLYSKMIVGNHKDRSRS

## FIGURE 168

GCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGCCCTCA  $\texttt{CGTCTCCTCCAGGG} \underline{\textbf{ATG}} \texttt{GCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCAC}$ ACCTGGCAGGCCCAGGCTGTTCCCACCATCCTGCCCCTGGGCCTGGCTCCAGACACCTTTGA CGATACCTATGTGGGTTGTGCAGAGGAGATGGAGGAGAAGGCAGCCCCCTGCTAAAGGAGG AAATGGCCCACCATGCCTGCTGCGGGAATCCTGGGAGGCAGCCCAGGAGACCTGGGAGGAC AAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGT CTACACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAG GCTCCCGGGAGCTCTACATGAGGCACTTTCCCTTCAAGGCCCTGCATTTCTACCTGATCCGG GCCCTGCAGCTGCGAGGCAGTGGGGGGCTGCAGCAGGGGACCTGGGGAGGTGGTTTCCG AGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGACTCTGTCCGCTTGGGCCAGT TTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCCACAGATTTGGGGAGAAGAGGCGGGGCTGT GTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTCTGCCCCCCTG  ${\tt GAAGACTCTGCTCTTGGCCCCTGGAGAGTTCCAGCTCTCAGGGGTTGGGCCC} {\color{blue}{\textbf{TGA}}} {\color{blue}{\textbf{A}}} {\color{blue}{\textbf{A}}} {\color{blue}{\textbf{A}}} {\color{blue}{\textbf{C}}} {$ ACATCTGCCACTTAGGAGCCCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAG CAGCCTTGAGAAGCAAGAACATGGTTCCGGACCCAGCCCTAGCAGCCTTCTCCCCAACCAGG ATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACTCTGCTATGTGATGGGGACTTCCT TGGAGTTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGGA

## FIGURE 169

MALAALMIALGSLGLHTWQAQAVPTILPLGLAPDTFDDTYVGCAEEMEEKAAPLLKEEMAHH ALLRESWEAAQETWEDKRRGLTLPPGFKAQNGIAIMVYTNSSNTLYWELNQAVRTGGGSREL YMRHFPFKALHFYLIRALQLLRGSGGCSRGPGEVVFRGVGSLRFEPKRLGDSVRLGQFASSS LDKAVAHRFGEKRRGCVSAPGVQLGSQSEGASSLPPWKTLLLAPGEFQLSGVGP

## FIGURE 170

 $\tt GTGGCTTCATTTCAGTGGCTGACTTCCAGAGAGCAAT \underline{\textbf{ATG}} \texttt{GCTGGTTCCCCAACATGCCTCA}$  $\verb|CCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTG|\\$ GTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTC TATTGTCTGGACCTTCAACACACCCCTCTTGTCACCATACAGCCAGAAGGGGGCACTATCA TAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAG CTCAGCAAACTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACT CCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAG TCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATG GAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCTGGGGCCAAGCAGCCAATGAGTC CCATAATGGGTCCATCCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCT GCGTTGCCAGGAACCCTGTCAGCAGAAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGT GAAGGTGCTGCTGATGACCCAGATTCCTCCATGGTCCTCCTGTGTCTCCTGTTGGTGCCCCT AGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGGAAACTCCTAACATATGCCCCCAT TCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGA TCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGATGGAAAATCCCCACTCAC  ${\tt TGCTCACGATGCCAGACACCAAGGCTATTTGCCTATGAGAATGTTATC}$ CACTCCCCTAAGTCTCTGCTCA

## FIGURE 171

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVT
IQPEGGTIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHV
YEHLSKPKVTMGLQSNKNGTCVTNLTCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRW
GESDMTFICVARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLW
FLKRERQEEYIEEKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIP
KKMENPHSLLTMPDTPRLFAYENVI

### FIGURE 172

CTGGTTCCCCAACATGCCTCACCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCC TCTGGACCCGTGAAAGAGCTGGTCGGTTCCGTTGGTGGGGCCCGTGACTTTCCCCCTGAAGTC CAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTCACCATAC AGCCAGAAGGGGGCACTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCA GATGGAGGCTACTCCCTGAAGCTCAGCAAACTGAAGAAGAATGACTCAGGGATCTACTATGT GGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACG AGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTG ACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCT GGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAG AAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAAACTTCTCAAGCCCC ATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCTCCATGGTCCTCCT GTGTCTCCTGTTGGTGCCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTC ACTCCTAACATATGCCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAA TAGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAA AGATGGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAG 

### FIGURE 173

## FIGURE 174

MKMLLLCLGLTLVCVHAEEASSTGRNFNVEKINGEWHTIILASDKREKIEEHGNFRLFLEQ IHVLENSLVLKVHTVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLI NEKDGETFQLMGLYGREPDLSSDIKERFAQLCEEHGILRENIIDLSNANRCLQARE

## FIGURE 175

# FIGURE 176

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMA
IPATTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNA
NCEFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRL
IHFSVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

## FIGURE 177

# FIGURE 178

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNLQVAKLNPPPEALAAKLEV KHCTDQISFKKRLSLKKSWWK

### FIGURE 179

# FIGURE 180

 ${\tt MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCK} \\ {\tt YKSSQKQHSPVPEKAIPLITPGSATTC}$ 

## FIGURE 181

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGG**ATG**TCGCTGCTGAGCCTGCCC TGGCTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTC CTGGCTACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACTGCCGCCGGCTCC AGTGTTTCCCACAGCCCCCAAAACGGAACTGGTTTTGGGGTCACCTGGGCCTGATCACTCCT ACAGAGGAGGGCTTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGT ATGGCTGGGTCCCATCATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCA CCAATGCCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCCTGAAGCCCTGG CTGGGAGAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGAC GCCCGCCTTCCATTTCAACATCCTGAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACA TCATGCTTGACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAG CACATCAGCCTCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACAGCCATTG TCAGGAGAGGCCCAGTGAATATTTGCCACCATCTTGGAGCTCAGTGCCCTTGTAGAGAAAA GAAGCCAGCATATCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGCGGCGCGC TTCCACAGGGCCTGCCGCCTGGTGCATGACTTCACAGACGCTGTCATCCGGGAGCGGCGTCG CACCCTCCCCACTCAGGGTATTGATGATTTTTTTCAAAGACAAAGCCAAGTCCAAGACTTTGG ATTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTCAGATGAGGAT ATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCCTCTC CTGGGTCCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGC AAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCCC TTCCTGACCATGTGCGTGAAGGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATCTCCCG ATGCTGCACCCAGGACATTGTTCTCCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTGCC TCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGAC  ${\tt CCCTTCCGCTTTGACCCAGAGAGCAGCAAGGGGGGGGGTCACCTCTGGCTTTTATTCCTTTCTC}$ CGCAGGGCCCAGGAACTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCCTGG CGTTGATGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGCAGGAAGCTGGAA  ${\tt TTGATCATGCGCGCGGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCA}$ GTGACTTTCTGACCCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAAA

## FIGURE 182

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPQPPKRNWFWG
HLGLITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLF
IRFLKPWLGEGILLSGGDKWSRHRRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGS
SRLDMFEHISLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYY
LSHDGRRFHRACRLVHDFTDAVIRERRRTLPTQGIDDFFKDKAKSKTLDFIDVLLLSKDEDG
KALSDEDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCRQEVQELLKDRDPKEIEW
DDLAQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVW
PDPEVYDPFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHT
EPRRKLELIMRAEGGLWLRVEPLNVGLQ

## FIGURE 183

# FIGURE 184

MYKLASCCLLFTGFLNPLLSLPLLDSREISFQLSAPHEDARLTPEELERASLLQILPEMLGA ERGDILRKADSSTNIFNPRGNLRKFQDFSGQDPNILLSHLLARIWKPYKKRETPDCFWKYCV

## FIGURE 185

# FIGURE 186

MPSPGTVCSLLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDG GQAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

## FIGURE 187

CGGCCACAGCTGGCATGCTCTGCCTGATCGCCATCCTGCTGTATGTCCTCGTCCAGTACCTC GTGAACCCCGGGGTGCTCCGCACGGACCCCAGATGTCAAGAAT**ATG**AACACGTGGCTGCTGT TCCTCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTG CTCCTGCTGGACTTTCTTGGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCT GAGTATGTCCCCCACCCTAAGCCCCCGATCCCCCCAAGGCTGGGTGGTCAGAGCTGCTCATC TTACACCTCTACTTGAGTATGTCCCTAACCCTGAGCCCCCCACGCCTGGGGCCAGAGTCTTT GTCCCCCGTGTGCGCATGTGTTCAGGGTCAGCCTCTCCCAGAAGTGAGATCATGGACAAAAA GGGCAAATCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCCAGCAAGAAGCTGAAC TCACGCCGAGACCTGCAGGAGTGCTGCCAGGTGCT**TGA**AGTAACAAGTTTAAAATGTTCAGA GACAATGGAATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTT AACAACTGAAGCGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGTC AGTCATGTTGCTGAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAAACCTGTGTTGGAAGT AACAACACCTCCTGCTCCTGGCACCAGCCGTTTTGGTCATGGTGGGCCAGCTGCAAAGCG TCTTCCATTCTCTGGGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTTTCTGTGGAC ACGGGCAGCAGAGTGTGTCCAGGCCAGCCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCA ACCCCTGGTCAGGGCAGAGGGAGTTGGGTGGGTCAGGCTCTGGGCTCACCTCCATCTCCAGA GCATCCCCTGCCTGCAGTTGTGGCAAGAACGCCCAGCTCAGAATGAACACCCCCACCAAGA GCCTCCTTGTTCATAACCACAGGTTACCCTACAAACCACTGTCCCCACACAACCCTGGGGAT GTTTTAAAACACACCTCTAACGCATATCTTACAGTCACTGTTGTCTTGCCTGAGGGTTGA ATTTTTTTAATGAAAGTGCAATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAA 

# FIGURE 188

MNTWLLFLPLFPVQVQTLIVVIIGMLVLLLDFLGLVHLGQLLIFHIYLSMSPTLSPRSPQGW VVRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQ AQQEAELTPRPAGVVPGA

## FIGURE 189

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACCATGGCCAAG ATGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCT ATCACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGG TGCCCAAGCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCCTG CCGGTTCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAAACTGTGGAAG GGACTACTGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAA GCGGTTGATGGAGAAGGCTTCCCTCCCCTCCCTTGGGGCTTTGTGGCAAAAATCCTA TGGTTATCCCTGGGAACGCAGATCACCTACATCGGACTTCAATTCATCAGCTTCCTCCTGCT ACTAACAGACTTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAACTGAGCGCCTTTGCTG CTGTTTCCTCTGTCCTGTCAGGTCTCCTGGGGATGGTGGCCCACATGATGTATTCACAAGTC TTCCAAGCGACTGTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGAATTATGGCTG GGCCTTCTACATGGCCTGGCTCTCCTTCACCTGCTGCATGGCGTCGGCTGTCACCACCTTCA ACACGTACACCAGGATGGTGCTGGAGTTCAAGTGCAAGCA**TAG**TAAGAGCTTCAAGGAAAAC CCGAACTGCCTACCACATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCCAC CGTGGGTCCTTTGACCAGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGG AAAGAAGCAGTTAGGTCATCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGA GTAGGCTTGAGCCCTACCTTACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCG TCTCTTGAGCATGGTTTTTAGAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTC CTAAGGGATTCCTGGGTGCCACTGCTCTTTTCCTCTACAGCTCCATCTTGTTTCACCCAC CCCACATCTCACACATCCAGAATTCCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGC TAAACCATGGAGATAAAAAGAAGAGTAAAATACACTTCCCGACCTTAAGGATCTGAAA

# FIGURE 190

MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWFVGTQKVPKPLCEKGLAAKCFDMP VSLDGDTNTSTQEVVQYNWETGDDRFSFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKR GEKGLLEFATLQGPCHPTLRFGGKRLMEKASLPSPPLGLCGKNPMVIPGNADHLHRTSIHQL PPATNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSDCQLGSRRLETTCLE LWLGLLHGLALLHLLHGVGCHHLQHVHQDGAGVQVQA

# FIGURE 191

AACTGGAAGGAAAGAAAGGTCAGCTTTGGCCCAG**ATG**TGGTTACCCCTTGGTCTCCTG TCTTTATGTCTTTCTCCTCTTCCTATTCTGTCATCTCCCTCACTTAAGTCTCAGGCCTGTCA GCAGCTCCTGTGGACATTGCCATCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATG TTATGGATGTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGT GATCTCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTGAAATCTGGCATG GTCTGTTCTCTTATTGTCAACCTCAGCACAACAGGCTGGCGCCAATGGCATTACAGAGAAAG CAATCTGTGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGT AGCCACCTCCCTGTCAGCCAGTATTAACATGTCCCCTTCCCCCTGCCCGCCGTAGATTCAG GACATTCGCCCCTGTGTGCCACCAAACCAGGACTTTCCCCCTTGGCTTGGCATCCCTGGCTCT CTCCTGGTACCCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTTACT ATGGCGATGGCCATGATGTTACAATCCCACTTGCCTGAATAATCAAGTGGGAAGGGGAAGCA GAGGGAAATGGGGCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCTGAGGAAAAACCAAA TGTTGAAGGGGCACAAGAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAA ATAGAACAGTCTGCTGGGAGTCAGACCTGGAATTCTGATTCCAAACTCTTTATTACTTTGGG AAGTCACTCAGCCTCCCCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGG AACCAAGGAAACTAACAATGTAGGTTACTAGTGAATACCCCAATGGTTTCTCCAATTATGCC CATGCCACCAAAACAATAAAACAAAATTCTCTAACACTGAAA

# FIGURE 192

 $\verb|MWLPLGLLSLCLSPLPILSSPSLKSQACQQLLWTLPSPLVAFRANRTTYVMDVSTNQGSGME| \\ | HRNHLCFCDLYDRATSPPLKCSLL| \\$ 

## FIGURE 193

CCGCC ATG GCAGGCATCAAAGCTTTGATTAGTTTGTCCTTTGGAGGAGCAATCGGACTGATGTTTTTGATGCTTGGATGTGCCCTTCCAATATACAACAAATACTGGCCCCTCTTTGTTCTATTTTTTTACATCCTTTCACCTATTCC ATACTGCATAGCAAGAAGATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACTTGCCATCTTTC TTACAACGGGCATTGTCGTGTCAGCTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGATTGAGTGGGGA GCTTGTGCACTTGTTCTCACAGGAAACACAGTCATCTTTGCAACTATACTAGGCTTTTTCTTGGTCTTTGGAAG  ${\tt CAATGACGACTTCAGCTGGCAGCAGTGG} \underline{{\tt TGA}}{\tt AAAGAAATTACTGAACTATTGTCAAATGGACTTCCTGTCATTT}$ GTTGGCCATTCACGCACACAGGAGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTA  ${\tt GGTGCTCCCTTCTCACTTTTATTGTAAGCATACTATTTTCACAGAGACTTGCTGAAGGATTAAAAGGATTTTCT}$ CTTTTGGAAAAGCTTGACTGATTTCACACTTATCTATAGTATGCTTTTTTGTGGTGTCCTGCTGAATTTAAATAT TTATGTGTTTTTCCTGTTAGGTTGATTTTTTTTGGAATCAATATGCAATGTTAAACACTTTTTTAATGTAATCA TTTGCATTGGTTAGGAATTCAGAATTCCGCCGGCTCTATTACTGGTCAAGTACATCTTTTCTCTTAAAATTATT CAGACATACAGACGGTTGGCATACGTTATAGACTGTATACTCAGTGCAAATATAGCTGCATTTATACCTCAGAG GGGCCAAGTGTTAATGCCCATGCCCTCCGTTAAGGGTTGTTGGTTTTACTGGTAGACAGATGTTTTGTGGATTG AAAATTATTTATGGAATTGCTACAGAGGAGTGCTTTTCTTCTCAATTGTTAGAAGAATTTATGTTAAACTTTA GGGAAGAATGACATTGAAATTCCAGTTTTTGAATCCTGTTTCTATTTATAAGTGAAATTTGTGATCTCCTATC AACCTTTCATGTTTTACCCTGTTAAAATGGACATACATGGAACCACTACTGATGAGGGACAGTTGTATGTTTGC CAGAGTGCCCCCTCCCCTGCAAGGCCTTGCCATGATTAACAAGTAACTTGTTAGTCTTACAGATAATTCATGCA TTAACAGTTTAAGATTTAGACCATGGTAATAGTAGTTCTTATTCTCTAAGGTTATATCATATGTAATTTAAAAG TATTTTTAAGACAAGTTTCCTGTATACCTCTGAACTGTTTTGATTTTGAGTTCATCATGATAGATCTGCTGTTT CCTTATAAAAGGCATTTGTTGTGTGAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACAT ACCTGACCAAAAAATTCCCAGTAACCAGGCATGATCAATTTATAGTGGTCGTTTACATCTAATAATTATCAGGA TTTATGAAGTTTATTTCTCAAGAAAATGGGAATAAATTTGGGATTTGTTCAGCTTTTTTACTAAAGATGCCTAA AGCCACAGGTTTTATTGCCTAACTTAAGCCATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCG GCGTGTGGCTGGAGCCTTCCCACTGGAGGCTGAAAGTGGCTTGTGGTATTATAATGTTCAGATTTCAAGAGGAA GGTGCAGGTACACATGAGTTAGAGAGCTGGTGAGACAGTTGGGAACTCTTTGTGCTTGTGATCTACTGGACTTT TTTTTTGCAGGAAGTGCATTCTCTGGTCCTTCCCTATTTTCTGTTCTGGATGTCAGTGCAGTGCACTGCTACTG TTTTATCCACTTGGCCACAGACTTTTTCTAACAGCTGCGTATTATTTCTATATACTAATTGCATTGGCAGCATT GTGTCTTTGACCTTGTATACTAGCTTGACATAGTGCTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAAT AGATTTTAAATATCTATTTTAAAAAAAAAA

# FIGURE 194

MAGIKALISLSFGGAIGLMFLMLGCALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTDAM SNACKELAIFLTTGIVVSAFGLPIVFARAHLIEWGACALVLTGNTVIFATILGFFLVFGSND DFSWQQW

### FIGURE 195

## FIGURE 196

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPH NLSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNHICSVQGDAFQKLRRVKELTLSSNQ ITQLPNTTFRPMPNLRSVDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRIFQDCRS LKFLDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRLISLHSLCLRRNKVAIV VSSLDWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKSLTSIT LAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDAVYAFHLCEDGAEPTSG HLLSAVTNRSDLGPPASSATTLADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMA LIFSFLIVVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQMAAMSAQEYYVDYKPNH IEGALVIINEYGSCTCHQQPARECEV

## FIGURE 197

GTGCAAGGAGCCGAGGCGAGATCGGCCGTCCTGGGCCGGGTCCTGCTGTGCTGCAGCTCTGC
GCACTGACCCAGGCGGTCTCCAAACTCTGGGTCCCCAACACGGACTTCGACGTCGCAGCCAA
CTGGAGCCAGAACCGGACCCCGTGCGCCGGCGGCGCGCGTTGAGTTCCCGGCGGACAAGATGG
TGTCAGTCCTGGTGCAAGAAGGTCACGCCGTCTCAGACATGCTCCTGCCGCTGGATGGGGAA
CTCGTCCTGGCTTCAGGAGCCGGATTCGGCGTCTCAGACGTGGGCTCGCACCTGGACTGTGG
CGCGGGCGAACCTGCCGTCTTCCGCGACTCTGACCGCTTCTCCTGGCATGACCCGCACCTGT
GGCGCTCTGGGGACGAGGCACCTGGCCTCTTCTTCGTGGACGCCGAGCGCGTGCCCTGCCGC
CACGACGACGTCTTCTTTCCGCCTAGTGCCTCCTTCCGCGTGGGGCTCGGCCCTGGCGCTAG
CCCCGTGCGTGTCCGCAGCATCTCGGCTCTGGGCCGGACGACGACGACGACGTGG
CTGTTTTCCTGGCGTCCCGCGCGGGCCCCTACGCTTCCACGGGCCGGGCGCGCCTCTGAGCGTG
GGCCCCGAGGACTGCGCGGGCCGCTCTGCGCCTTCCACGGGCCGGAGGCGCAGCCGTG
GGCCCCGAGGACTGCGCGGACCCTTCCACGCTTCCACGGGCCGGAGGCGCAGCCGTG
GATCTGCGCGGCCCTGCTCCAGCCCCT

# FIGURE 198

MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLVQE GHAVSDMLLPLDGELVLASGAGFGVSDVGSHLDCGAGEPAVFRDSDRFSWHDPHLWRSGDEA PGLFFVDAERVPCRHDDVFFPPSASFRVGLGPGASPVRVRSISALGRTFTRDEDLAVFLASR AGRLRFHGPGALSVGPEDCADPSGCVCGNAEAQPWICAALLQP

## FIGURE 199

# FIGURE 200

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIP FARDAVKKCFAVCLA

## FIGURE 201

CCGAGGGAGTCTCCTCCAGACCTCCCTCCCGTTGCTCCAAACTAATACGGACTGAACGGATCGCTGCGAGGGT AACTGATCAAGTACTTTGAAAAATGACTTCGAAATTTATCTTGGTGTCCTTCATACTTGCTGCACTGAGTCTTTC AACCACCTTTTCTCTCCAACTAGACCAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACT TATATAAAGTTCCAACGCCCCATTTTCATTATATTATGAAATATGGTGTTCACGTGAAGCAAGTTACTAATGTT TTTATTACAAAAACCTACCCTAACCATTATACTTTGGTAACTGGCCTCTTTGCAGAGAATCATGGGATTGTTGC AAATGATATGTTTGATCCTATTCGGAACAAATCTTTCTCCTTGGATCACATGAATATTTATGATTCCAAGTTTT GGGAAGAGCGACACCAATATGGATCACAAACCAGAGGGCAGGACATACTAGTGGTGCAGCCATGTGGCCCGGA ACAGATGTAAAAATACATAAGCGCTTTCCTACTCATTACATGCCTTACAATGAGTCAGTTTCATTTGAAGATAG AGTTGCCAAAATTGTTGAATGGTTTACGTCAAAAGAGCCCATAAATCTTGGTCTTCTCTATTGGGAAGACCCTG ATGACATGGGCCACCATTTGGGACCTGACAGTCCGCTCATGGGGCCTGTCATTTCAGATATTGACAAGAAGTTA GGATATCTCATACAAATGCTGAAAAAGGCAAAGTTGTGGAACACTCTGAACCTAATCATCACAAGTGATCATGG AATGACGCAGTGCTCTGAGGAAAGGTTAATAGAACTTGACCAGTACCTGGATAAAGACCACTATACCCTGATTG ATCAATCTCCAGTAGCAGCCATCTTGCCAAAAGAAGGTAAATTTGATGAAGTCTATGAAGCACTAACTCACGCT CATCCTAATCTTACTGTTTACAAAAAAGAAGACGTTCCAGAAAGGTGGCATTACAAATACAACAGTCGAATTCA ACCAATCATAGCAGTGGCTGATGAAGGGTGGCACATTTTACAGAATAAGTCAGATGACTTTCTGTTAGGCAACC TCAAAAGAAGCCATGAACTCCACAGATTTGTACCCACTACTATGCCACCTCCTCAATATCACTGCCATGCCACA CAATGGATCATTCTGGAATGTCCAGGATCTGCTCAATTCAGCAATGCCAAGGGTGGTCCCTTATACACAGAGTA  $\verb|CTATACTCCTCCCTGGTAGTGTTAAACCAGCAGAATATGACCAAGAGGGGTCATACCCTTATTTCATAGGGGTC| \\$ TCTCTTGGCAGCATTATAGTGATTGTATTTTTGTAATTTCATTAAGCATTTAATTCACAGTCAAATACCTGC CTTACAAGATATGCATGCTGAAATAGCTCAACCATTATTACAAGCCTAATGTTACTTTGAAGTGGATTTGCATA TTGAAGTGGAGATTCCATAATTATGTCAGTGTTTAAAGGTTTCAAATTCTGGGAAACCAGTTCCAAACATCTGC ATCCTGCTTTATTTGGACTTGGCGCAGATAATGTATATTTTAGCAACTTTGCACTATGTAAAGTACCTTATAT ATTGCACTTTAAATTTCTCTCCTGATGGGTACTTTAATTTGAAATGCACTTTATGGACAGTTATGTCTTATAAC TTGATTGAAAATGACAACTTTTTGCACCCATGTCACAGAATACTTGTTACGCATTGTTCAAACTGAAGGAAATT TCTAATAATCCCGAATAATGAACATAGAAATCTATCTCCATAAATTGAGAGAAGAAGAAGGTGATAAGTGTTGA AAATTAAATGTGATAACCTTTGAACCTTGAATTTTGGAGATGTATTCCCAACAGCAGAATGCAACTGTGGGCAT ATTCGTTCTAAATATTGTTTCTGTCATAAAATTATTGTGATTTCCTGATGAGTCATATTACTGTGATTTTCA TAATAATGAAGACACCATGAATATACTTTTCTTCTATATAGTTCAGCAATGGCCTGAATAGAAGCAACCAGGCA AAATCAAATTGGATAAAAAAAAAAAAAAAAAAAAA

## FIGURE 202

MTSKFILVSFILAALSLSTTFSLQLDQQKVLLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVK
QVTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEEATPIW
ITNQRAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLY
WEDPDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEER
LIELDQYLDKDHYTLIDQSPVAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYN
SRIQPIIAVADEGWHILQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTD
LYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPYF
IGVSLGSIIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

#### Signal Peptide:

amino acids 1-22

#### Transmembrane Domain:

amino acids 429-452

## N-glycosylation sites:

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372, 382-385, 389-392

#### Somatomedin B Domain:

amino acids 69-85

#### Sulfatase protein Region:

amino acids 212-241

## FIGURE 203

GGATTTTTGTGATCCGCGATTCGCTCCCACGGGCGGACCTTTGTAACTGCGGGAGGCCCAG GACAGGCCCACCCTGCGGGGCGGGGGGGGCAGCCGGGGTGAGGAGGTGAAGAAACCAAGACGC AGAGAGGCCAAGCCCCTTGCCTTGGGTCACACAGCCAAAGGAGGCAGAGCCAGAACTCACAA  $\tt CCAGATCCAGAGGCAACAGGGAC\underline{ATG} GCCACCTGGGACGAAAAGGCAGTCACCCGCAGGGCC$ AAGGTGGCTCCCGCTGAGAGGATGAGCAAGTTCTTAAGGCACTTCACGGTCGTGGGAGACGA AGCAGCCACCACCCACACCAGTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCTGACGTTGCC CCTGCCCTGGCCCCGCACCCAGGGCCCCCTTGACTTCAGGGGCATGTTGAGGAAACTGTT CAGCTCCCACAGGTTTCAGGTCATCATCTGCTTGGTGGTTCTTGGATGCCCTCCTGGTGC TTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCGACAAGAATAACTATGCTGCCATG ATTTGTCTTCCGCCTGAGTTCTTTCACCACAAGTTTGAGATCCTGGATGCCCGTCGTGGTGG TGGTCTCATTCATCCTGGACATTGTCCTCCTGTTCCAGGAGCACCAGTTTGAGGCTCTGGGC CTGCTGATTCTGCTCCGGCTGTGGCCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGT TAAGACACGTTCAGAACGGCAACTCTTAAGGTTAAAACAGATGAATGTACAATTGGCCGCCA AGATTCAACACCTTGAGTTCAGCTGCTCTGAGAAGCCCCTGGAC $\underline{\mathbf{TGA}}$ TGAGTTTGCTGTATC AACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTACTCT CAGGCTGGCATGTTCACTGGGCTGGTTTACGACAGAGAACCTGACAGTCACTGGCCAGTTA TCACTTCAGATTACAAATCACAGAGCATCTGCCTGTTTTCAATCACAAGAGAACAAAACC AAAATCTATAAAGATATTCTGAAAATATGACAGAATTTGACAAATAAAAGCATAAACGTGTA 

# FIGURE 204

MATWDEKAVTRRAKVAPAERMSKFLRHFTVVGDDYHAWNINYKKWENEEEEEEEQPPPTPV SGEEGRAAAPDVAPAPGPAPRAPLDFRGMLRKLFSSHRFQVIIICLVVLDALLVLAELILDL KIIQPDKNNYAAMVFHYMSITILVFFMMEIIFKLFVFRLSSFTTSLRSWMPVVVVVSFILDI VLLFQEHQFEALGLLILLRLWRVARIINGIIISVKTRSERQLLRLKQMNVQLAAKIQHLEFS CSEKPLD

## FIGURE 205

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCAGTGGAGCACCCAGCAGGCCGCCAAC<u>**AT**</u>  $\mathbf{\underline{G}}$ CTCTGTCTGTGCCTGTACGTGCCGGTCATCGGGGAAGCCCAGACCGAGTTCCAGTACTT $\overline{\mathrm{TG}}$ ĀGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCCATTTTCAAGCTCAGTGTCTTCATCCCCTCC CAGGAATTCTCCACCTACCGCCAGTGGAAGCAGAAAATTGTACAAGCTGGAGATAAGGACCT TGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGA GGCTGGTGTTTAAGATTTTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATG CATGGATAAAAACGGCACGATGACCATCGACTGGAACGAGTGGAGAGACTACCACCTCCTCC ACCCCGTGGAAAACATCCCCGAGATCATCCTCTACTGGAAGCATTCCACGATCTTTGATGTG GGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGAGGAGGCAGACGGGGATGTGGTG GAGACACCTGGTGGCAGGAGGTGGGGCAGGGCCGTATCCAGAACCTGCACGGCCCCCTGG ACAGGCTCAAGGTGCTCATGCAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGT GGCTTCACTCAGATGATTCGAGAAGGAGGGGCCAGGTCACTCTGGCGGGGCAATGGCATCAA TGGTGCTCTGAGCTGGACCCTGTCAGGATGGCCCCACCTCAGAACCAAACTCACTG TCCCCACTGTGGCATGAGGCAGTGGAGCACCATGTTTGAGGGCGAAGGGCAGAGCGTTTGT TCCCCACTGTGGGAAGGAAAAGGTGTTGGGAGCCCTTAATTATGGGAACCAAAACGGGGTTTGT TTTTGTCCAGAAGGACAAGCCGGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG GAGCAGGAGCTTGGCTGACTCCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCCTGTCCAACC CCAGCAGGGGCGCAGCGGGACCAGCCCCACATTCCACTTGTGTCACTGCTTGGAACCTATTT CTGTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAGAAGGCAGCAGCCCTG GCTCCTTTCCTTTGGCAGGTTGGGGAAGGGCTTGCCCCAGCCTTAGGATTTCAGGGTTTGA CTGGGGGCGTGGAGAGAGAGGGAGGAACCTCAATAACCTTGAAGGTGGAATCCAGTTATTTC CTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAGTGAGGTGCCTCT 

# FIGURE 206

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKD LDGQLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILK SMDKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMW WRHLVAGGGAGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGI NVLKIAPESAIKFMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMAL RKTGQYSGMLDCARRILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNS ADPGVFVLLACGTMSSTCGQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFG LYRGLAPNFMKVIPAVSISYVVYENLKITLGVQSR

#### Important features:

#### Signal peptide:

amino acids 1-16

#### Transmembrane domain:

amino acids 284-304, 339-360, 376-394

## Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

#### N-glycosylation site.

amino acids 129-133, 169-173

## Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

## FIGURE 207

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCAT CAATTGCACTCATCATTGGCTTTGGTATTTCAGGGAGACACTCCATCACAGTCACTACTGTC GCCTCAGCTGGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAA ACTTTCTGATATCGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCA AAGAAGGCAAAGATGAGCTGTCGGAGCAGGATGAAATGTTCAGAGGCCGGACAGCAGTGTTT GCTGATCAAGTGATAGTTGGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGC TGGCACCTACAAATGTTATATCATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATA AAACTGGAGCCTTCAGCATGCCGGAAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTG CGGTGTGAGGCTCCCCGATGGTTCCCCCAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCA GGGAGCCAACTTCTCGGAAGTCTCCAATACCAGCTTTGAGCTGAACTCTGAGAATGTGACCA TGAAGGTTGTGTCTGTGCTCTACAATGTTACGATCAACAACACATACTCCTGTATGATTGAA AATGACATTGCCAAAGCAACAGGGGATATCAAAGTGACAGAATCGGAGATCAAAAGGCGGAG GGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAATGTGCCTTGGCCACAAAAAAG CATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTCACCACCAGATATGACCTAG TTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTGAGCAAACAAGAGCA GACATATTAGAAGTTGGGAAAATAATTCATGTGAACTAGACAAGTGTGTTAAGAGTGATAAG GGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGCTG TAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCA CAAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGG GGCGGCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCT TGGCTTCTCTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAA ACAGAGCAGTCGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTTAAACAAAAAA 

# FIGURE 208

MASLGQILFWSIISIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDI KLSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTD AGTYKCYIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVD QGANFSEVSNTSFELNSENVTMKVVSVLYNVTINNTYSCMIENDIAKATGDIKVTESEIKRR SHLQLLNSKASLCVSSFFAISWALLPLSPYLMLK

## FIGURE 209

GAATTTGTAGAAGACAGCGGCGTTGCC**ATG**GCGGCGTCTCTGGGGCAGGTGTTGGCTCTGGT GCTGGTGGCCGCTCTGTGGGGTGGCACGCAGCCGCTGCTGAAGCGGGCCTCCGCCGGCCTGC AGCGGGTTCATGAGCCGACCTGGGCCCAGCAGTTGCTACAGGAGATGAAGACCCTCTTCTTG AATACTGAGTACCTGATGCCCTTTCTCCTCAACCAGTGTGGATCCCTTCTCTATTACCTCAC CTTGGCATCGACAGATCTGACCCTGGCTGTGCCCATCTGTAACTCTCTGGCTATCATCTTCA CACTGATTGTTGGGAAGGCCCTTGGAGAAGATATTGGTGGAAAACGTAAGTTAGACTACTGC CTCCCCAGAGTGGGTGAGGACACGGCCTTTTCCCATCCTGCCCTTTCCTCTGCAGCTGTTTT GCTTCCTTGTGGCCATCAGAGTTCCCTTCCCCTGGACAGTCTGGAGAAAGACAGAGGCTGGG GTTTGGGAT**TGA**AGACCAGACCCCATCTGAGCCCTTCCTCCAGCCCTGTACCAGCTCCTACT GGCATGGCTGAGCTCAGACCCTCCTGATTTCTGCCTATTATCCCAGGAGCAGTTGCTGGCAT GGTGCTCACCGTGATAGGAATTTCACTCTGCATCACAAGCTCAGTGAGTAAGACCCAGGGGC AACAGTCTACCCTTTGAGTGGGCCGAACCCACTTCCAGCTCTGCTGCCTCCAGGAAGCCCCT GGGCCATGAAGTGCTGGCAGTGAGCGGATGGACCTAGCACTTCCCCTCTCTGGCCTTAGCTT CCTCCTCTCTTATGGGGATAACAGCTACCTCATGGATCACAATAAGAGAACAAGAGTGAAAG AGTTTTGTAACCTTCAAGTGCTGTTCAGCTGCGGGGATTTAGCACAGGAGACTCTACGCTCA CCCTCAGCAACCTTTCTGCCCCAGCAGCTCTCTTCCTGCTAACATCTCAGGCTCCCAGCCCA GCCACCATTACTGTGGCCTGATCTGGACTATCATGGTGGCAGGTTCCATGGACTGCAGAACT CCAGCTGCATGGAAAGGGCCAGCTGCAGACTTTGAGCCAGAAATGCAAACGGGAGGCCTCTG GGACTCAGTCAGAGCGCTTTGGCTGAATGAGGGGTGGAACCGAGGGAAGAAGGTGCGTCGGA AAATCCTCACTGCCAGCCCCTCTTAAACAGGTAGAGAGCTGTGAGCCCCAGCCCACCTGAC 

# FIGURE 210

MAASLGQVLALVLVAALWGGTQPLLKRASAGLQRVHEPTWAQQLLQEMKTLFLNTEYLMPFL LNQCGSLLYYLTLASTDLTLAVPICNSLAIIFTLIVGKALGEDIGGKRKLDYCECGTQLCGS RHTCVSSFPEPISPEWVRTRPFPILPFPLQLFCFLVAIRVPFPWTVWRKTEAGVWD

## FIGURE 211

CTTCTGTAGGACAGTCACCAGGCCAGATCCAGAAGCCTCTCTAGGCTCCAGCTTTCTCTGTG GAAGATGACAGCAATTATAGCAGGACCCTGCCAGGCTGTCGAAAAGATTCCGCAATAAAACT TTGCCAGTGGGAAGTACCTAGTGAAACGGCCTAAGATGCCACTTCTTCTCATGTCCCAGGCT TGAGGCCCTGTGGTCCCCATCCTTGGGAGAAGTCAGCTCCAGCACCATGAAGGGCATCCTCG TTGCTGGTATCACTGCAGTGCTTGTTGCAGCTGTAGAATCTCTGAGCTGCGTGCAGTGTAAT TCATGGGAAAAATCCTGTGTCAACAGCATTGCCTCTGAATGTCCCTCACATGCCAACACCAG CTGTATCAGCTCCTCAGCCAGCTCCTCTCTAGAGACACCAGTCAGATTATACCAGAATATGT TCTGCTCAGCGGAGAACTGCAGTGAGGAGACACACATTACAGCCTTCACTGTCCACGTGTCT GCTGAAGAACACTTTCATTTTGTAAGCCAGTGCTGCCAAGGAAAGGAATGCAGCAACACCAG CGATGCCCTGGACCCTCCCCTGAAGAACGTGTCCAGCAACGCAGAGTGCCCTGCTTGTTATG AATCTAATGGAACTTCCTGTCGTGGGAAGCCCTGGAAATGCTATGAAGAAGAACAGTGTGTC TTTCTAGTTGCAGAACTTAAGAATGACATTGAGTCTAAGAGTCTCGTGCTGAAAGGCTGTTC CAACGTCAGTAACGCCACCTGTCAGTTCCTGTCTGGTGAAAACAAGACTCTTGGAGGAGTCA TCCCACACGTGGGCTCCAAAGCTTCCCTCTACCTCTTGGCCCTTGCCAGCCTCCTTCTTCG GGGACTGCTGCCCTGAGGTCCTGGGGCTGCACTTTGCCCAGCACCCCATTTCTGCTTCTCTG AGGTCCAGAGCACCCCTGCGGTGCTGACACCCTCTTTCCCTGCTCTGCCCCGTTTAACTGC CCAGTAAGTGGGAGTCACAGGTCTCCAGGCAATGCCGACAGCTGCCTTGTTCTTCATTATTA 

# FIGURE 212

MKGILVAGITAVLVAAVESLSCVQCNSWEKSCVNSIASECPSHANTSCISSSASSSLETPVR LYQNMFCSAENCSEETHITAFTVHVSAEEHFHFVSQCCQGKECSNTSDALDPPLKNVSSNAE CPACYESNGTSCRGKPWKCYEEEQCVFLVAELKNDIESKSLVLKGCSNVSNATCQFLSGENK TLGGVIFRKFECANVNSLTPTSAPTTSHNVGSKASLYLLALASLLLRGLLP

# FIGURE 213

 $\mathsf{GGGCTTGCCTCACTGGCCACCCTCCCAACCCCAAGAGCCCCAGCCCC}$ ACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGCTTTGGGGGCCCCATGACCCGCAG CTACCGGAGCACCGCCCGGACTGGTCTTCCCCGGAAGACAAGGATAATCCTAGAGGACGAGA ATGATGCCATGGCCGACGCCGACCGCCTGGCTGGACCAGCGGCTGCCGAGCTCTTGGCCGCC ACGGTGTCCACCGGCTTTAGCCGGTCGTCCGCCATTAACGAGGAGGATGGGTCTTCAGAAGA GGGGGTTGTGATTAATGCCGGAAAGGATAGCACCAGCAGAGAGCTTCCCAGTGCGACTCCCA ATACAGCGGGGAGTTCCAGCACGAGGTTTATAGCCAATAGTCAGGAGCCTGAAATCAGGCTG ACTTCAAGCCTGCCGCGCTCCCCCGGGAGGTCTACTGAGGACCTGCCAGGCTCGCAGGCCAC CCTGAGCCAGTGGTCCACACCTGGGTCTACCCCGAGCCGGTGGCCGTCACCCTCACCCACAG CCATGCCATCTCCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCGTGGCACTGCCAC TGCAAGTCGGGCACCATGAGCCGGAGCCGGTCTGGGAAGCTGCACGGCCTTTCCGGGCCCCT TCGAGTTGGGGCCCTGAGCCACCCCCCCCCCCACGAGCCCTTGCACCTATCAACAATGTC CCTGCAACCGACTTCGGGAAGAGTGCCCCCTGGACACAAGTCTCTGTACTGACACCAACTGT GCCTCTCAGAGCACCACCAGTACCAGGACCACCACTACCCCCTTCCCCACCATCCACCTCAG AAGCAGTCCCAGCCTGCCACCCGCCAGCCCTGCCCAGCCTTTTTTGGAAACGGGTCA GGATTGGCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTCACAGAGATGCAACCAATA  ${\sf GACAGAAACCAGAGG}$   ${\sf TAA}$   ${\sf TGGCCACTTCATCCACATGAGGAGATGTCAGTATCTCAACCTCT}$ CTTGCCCTTTCAATCCTAGCACCCACTAGATATTTTTAGTACAGAAAAACAAAACTGGAAAA CACAA

## FIGURE 214

MVPAAGALLWVLLLNLGPRAAGAQGLTQTPTEMQRVSLRFGGPMTRSYRSTARTGLPRKTRI ILEDENDAMADADRLAGPAAAELLAATVSTGFSRSSAINEEDGSSEEGVVINAGKDSTSREL PSATPNTAGSSSTRFIANSQEPEIRLTSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWP SPSPTAMPSPEDLRLVLMPWGPWHCHCKSGTMSRSRSGKLHGLSGRLRVGALSQLRTEHKPC TYQQCPCNRLREECPLDTSLCTDTNCASQSTTSTRTTTTPFPTIHLRSSPSLPPASPCPALA FWKRVRIGLEDIWNSLSSVFTEMQPIDRNQR

## FIGURE 215

GCGGCGCTGGCGACCGCTCCCAGGGCGACCGTGAGCCGGTGTACCGCGACTGCGTACTGCGAAGAGCA GTCGGGACGACTGTAAGTATGAGTGTATGTGGGTCACCGTTGGGCTCTACCTCCAGGAAGGTCACAAAGTGCCT CAGTTCCATGGCAAGTGGCCCTTCTCCCGGTTCCTGTTCTTCAAGAGCCGGCATCGGCCGTGGCCTCGTTTCT CAATGGCCTGGCCAGCCTGGTGATGCTCTGCCGCTACCGCACCTTCGTGCCAGCCTCCTCCCCCATGTACCACA  $\tt CCTGTGTGGCCTTCGCCTGGGTGTCCCTCAATGCATGGTTCTGGTCCACAGTCTTCCACACCAGGGACACTGAC$  $\tt CTCACAGAGAAAATGGACTACTTCTGTGCCTCCACTGTCATCCTACACTCAATCTACCTGTGCTGCGTCAGGAC$  $\tt CCTACCTGAGCCTCATCCGCTTCGACTATGGCTACAACCTGGTGGCCAACGTGGCTATTGGCCTGGTCAACGTG$ GTGTGGTGGCTGGCCTGTGGAACCAGCGGCGGCTGCCTCACGTGCGCAAGTGCGTGGTGGTGGTCTT GCTGCTGCAGGGGCTGTCCCTGCTCGAGCTGCTTGACTTCCCACCGCTCTTCTGGGTCCTGGATGCCCATGCCA TCTGGCACATCAGCACCATCCCTGTCCACGTCCTCTTTTTCAGCTTTCTGGAAGATGACAGCCTGTACCTGCTG  $AAGGAATCAGAGGACAAGTTCAAGCTGGACT{\color{red}GAGACCTTGGAGCGAGTCTGCCCCAGTGGGGATCCTGCCCCC}\\$ GCCCTGCTGGCCTCCCTTCTCCCCTCAACCCTTGAGATGATTTTCTCTTTTCAACTTCTTGAACTTGGACATGA AGGATGTGGGCCCAGAATCATGTGGCCAGCCCACCCCTGTTGGCCCTCACCAGCCTTGGAGTCTGTTCTAGGG AAGGCCTCCCAGCATCTGGGACTCGAGAGTGGGCAGCCCCTCTACCTCCTGGAGCTGAACTGGGGTGGAACTGA GTGTGTTCTTAGCTCTACCGGGAGGACAGCTGCCTGTTTCCTCCCCACCAGCCTCCTCCCCACATCCCCAGCTG  $\tt CCTGGCTGGGTCCTGAAGCCCTCTGTCTACCTGGGAGACCAGGGACCACAGGCCTTAGGGATACAGGGGGTCCC$ GGTTCACGGCGATTCTCCCCATGGGATCTTGAGGGACCAAGCTGCTGGGATTGGGAAGGAGTTTCACCCTGACC GTTGCCCTAGCCAGGTTCCCAGGAGGCCTCACCATACTCCCTTTCAGGGCCAGGGCTCCAGCAAGCCCAGGGCA AGGATCCTGTGCTGCTGTTGAGAGCCTGCCACCGTGTGCGGGGTGTGGGCCAGGCTGAGTGCATAGG TGACAGGGCCGTGAGCATGGGCCTGGGTGTGTGAGCTCAGGCCTAGGTGCGCAGTGTGGAGACGGGTGTTGT GAGGGAATCCTGTCACCATCAATAATCACTTGTGGAGCGCCAGCTCTGCCCAAGACGCCACCTGGGCGGACAGC CAGGAGCTCTCCATGGCCAGGCTGCCTGTGTGCATGTTCCCTGTCTGGTGCCCCTTTGCCCGCCTCCTGCAAAC  $\tt CTCACAGGGTCCCCACACAACAGTGCCCTCCAGAAGCAGCCCCTCGGAGGCAGAGGAAAATGGGGATGGC$ TGGGGCTCTCTCCATCCTCTTTTCTCCTTGCCTTCGCATGGCTTGCCCTCCCAAAACCTCCATTCCCCT GCTGCCAGCCCCTTTGCCATAGCCTGATTTTGGGGAGGAGGAGGGGGCGATTTGAGGGAGAAGGGGAGAAAGCT ACACTATGCCTGTGCCCTGGTAAAGGTGACCCCTGCCATTTACCAGCAGCCCTGGCATGTTCCTGCCCCACAGG AATAGAATGGAGGGAGCTCCAGAAACTTTCCATCCCAAAGGCAGTCTCCGTGGTTGAAGCAGACTGGATTTTTG  $\tt CTCTGCCCCTGACCCCTTGTCCCTCTTTGAGGGAGGGGAGCTATGCTAGGACTCCAACCTCAGGGACTCGGGTG$ GCCTGCGCTAGCTTCTTTTGATACTGAAAACTTTTAAGGTGGGAGGGTGGCAAGGGATGTGCTTAATAAATCAA TTCCAAGCCTCAAAAAAAAAAAAAAAAAA

# FIGURE 216

MAGLAARLVLLAGAAALASGSQGDREPVYRDCVLQCEEQNCSGGALNHFRSRQPIYMSLAGW
TCRDDCKYECMWVTVGLYLQEGHKVPQFHGKWPFSRFLFFQEPASAVASFLNGLASLVMLCR
YRTFVPASSPMYHTCVAFAWVSLNAWFWSTVFHTRDTDLTEKMDYFCASTVILHSIYLCCVR
TVGLQHPAVVSAFRALLLLMLTVHVSYLSLIRFDYGYNLVANVAIGLVNVVWWLAWCLWNQR
RLPHVRKCVVVVLLLQGLSLLELLDFPPLFWVLDAHAIWHISTIPVHVLFFSFLEDDSLYLL
KESEDKFKLD

#### Important features:

### Signal peptide:

amino acids 1-20

### Transmembrane domains:

amino acids 105-123, 138-156, 169-185, 193-209, 221-240, 256-272

#### N-glycosylation site.

amino acids 40-44

#### N-myristoylation site.

amino acids 43-49

## CUB domain proteins profile.

amino acids 285-302

## Amiloride-sensitive sodium channels proteins.

amino acids 162-186

AND THE PROPERTY AND THE PROPERTY OF THE PROPE

### FIGURE 217

GGCCGCCTGGAATTGTGGGAGTTGTCTGCCACTCGGCTGCCGGAGGCCGAAGGTCCGTGA CTATGCTCCCCAGAGCCTGCCTTCATCTAGGATGGCTCCTCTGGGCATGCTGCTTGGGCTG CTGATGGCCGCCTGCTTCACCTTCTGCCTCAGTCATCAGAACCTGAAGGAGTTTGCCCTGAC TGGATGCCGAAGTCCTGGAGGTGTTCCACCCGACGCATGAGTGGCAGGCCCTTCAGCCAGGG ACTCCAATATGAGGACAAGTTCCGAAATAATTTGAAAGGCAAAAGGCTGGATATCAACACCA ACACCTACACATCTCAGGATCTCAAGAGTGCACTGGCAAAATTCAAGGAGGGGGCAGAGATG GAGAGTTCAAAGGAAGACAAGGCAAGGCAGGCTGAGGTAAAGCGGCTCTTCCGCCCCATTGA GGAACTGAAGAAGACTTTGATGAGCTGAATGTTGTCATTGAGACTGACATGCAGATCATGG TACGGCTGATCAACAAGTTCAATAGTTCCAGCTCCAGTTTGGAAGAGAAGATTGCTGCGCTC TTTGATCTTGAATATTATGTCCATCAGATGGACAATGCGCAGGACCTGCTTTCCTTTGGTGG TCTTCAAGTGGTGATCAATGGGCTGAACAGCACAGAGCCCCTCGTGAAGGAGTATGCTGCGT TTGTGCTGGGCGCTGCCTTTTCCAGCAACCCCAAGGTCCAGGTGGAGGCCATCGAAGGGGGA GCCCTGCAGAAGCTGCTGGTCATCCTGGCCACGGAGCAGCCGCTCACTGCAAAGAAGAAGGT CCTGTTTGCACTGTGCTCCCTGCTGCGCCACTTCCCCTATGCCCAGCGGCAGTTCCTGAAGC TCGGGGGGCTGCAGGTCCTGAGGACCCTGGTGCAGGAGAAGGGCACGGAGGTGCTCGCCGTG CGCGTGGTCACACTGCTCTACGACCTGGTCACGGAGAAGATGTTCGCCGAGGAGGAGGCTGA GCTGACCCAGGAGATGTCCCCAGAGAAGCTGCAGCAGTATCGCCAGGTACACCTCCTGCCAG GCCTGTGGGAACAGGGCTGGTGCGAGATCACGGCCCACCTCCTGGCGCTGCCCGAGCATGAT GCCCGTGAGAAGGTGCTGCAGACACTGGGCGTCCTCCTGACCACCTGCCGGGACCGCTACCG TCAGGACCCCCAGCTCGGCAGGACACTGGCCAGCCTGCAGGTACCAGGTGCTGGCCA GCCTGGAGCTGCAGGATGGTGAGGACGAGGGCTACTTCCAGGAGCTGCTGGGCTCTGTCAAC AGCTTGCTGAAGGAGCTGAGATGAGGCCCCACACCAGGACTGGACTGGGATGCCGCTAGTGA GGCTGAGGGTGCCAGCGTGGGTGGGCTTCTCAGGCAGGAGGACATCTTGGCAGTGCTGGCT 

### FIGURE 218

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPEKSSTKETERKETKAEEEL
DAEVLEVFHPTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGKRLDINTN
TYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMV
RLINKFNSSSSSLEEKIAALFDLEYYVHQMDNAQDLLSFGGLQVVINGLNSTEPLVKEYAAF
VLGAAFSSNPKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKL
GGLQVLRTLVQEKGTEVLAVRVVTLLYDLVTEKMFAEEEAELTQEMSPEKLQQYRQVHLLPG
LWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGRTLASLQAEYQVLAS
LELQDGEDEGYFQELLGSVNSLLKELR

#### Important features:

#### Signal peptide:

amino acids 1-29

#### Hypothetical YJL126w/YLR351c/yhcX family protein.

amino acids 364-373

### N-glycosylation site.

amino acids 193-197, 236-240

#### N-myristoylation site.

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

### Homologous region SLS1 protein.

amino acids 68-340

### FIGURE 219

TTCGGCTTCCGTAGAGGAAGTGGCGCGGACCTTCATTTGGGGTTTCGGTTCCCCCCCTTCCC CTTCCCCGGGGTCTGGGGGTGACATTGCACCGCGCCCCTCGTGGGGTCGCGTTGCCACCCCA CGCGGACTCCCCAGCTGGCGCGCCCCTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCCC CCGCCTTCGCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGT CGCAGGGGCATTTTTCTGGCTGGTCTCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGG TCCATGTGACCGACCGGTCAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCT GTCTCTGTCCTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGA TGAAGGGTTAGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCT ATGTTTCTGGTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCT GATGCACTTGGGCCAGGTGTGGTTGGGATCCATGGAGACTCACCCTATTACTTCCTGACTTC AGCCTTTCTGACAGCAGCCATTATCCTGCTCCATACCTTTTGGGGAGTTGTGTTCTTTGATG  ${\tt CCTGTGAGAGGGGGTACTGGGCTTTGGGCCTGGTGGTTGGGAGTCACCTACTGACATCG}$ GGACTGACATTCCTGAACCCCTGGTATGAGGCCAGCCTGCTGCCCATCTATGCAGTCACTGT TTCCATGGGGCTCTGGGCCTTCATCACAGCTGGAGGGTCCCTCCGAAGTATTCAGCGCAGCC CCCATGACTGAGCCCAGCCCCAGCCCGGGTCCATTGCCCACATTCTCTGTCTCCTTCTCGTC GGTCTACCCCACTACCTCCAGGGTTTTGCTTTGTCCTTTTGTGACCGTTAGTCTCTAAGCTT TACCAGGAGCAGCCTGGGTTCAGCCAGTCAGTGACTGGTGTGTTTGAATCTGCACTTATCCC CACCACCTGGGGACCCCCTTGTTGTGTCCAGGACTCCCCCTGTGTCAGTGCTCTGCTCTCAC CCTGCCCAAGACTCACCTCCCTTCCCCTCTGCAGGCCGACGCAGGAGGACAGTCGGGTGAT GGTGTATTCTGCCCTGCGCATCCCACCCGAGGACTGAGGGAACCTAGGGGGGACCCCTGGGC CTGGGGTGCCCTCTGATGTCCTCGCCCTGTATTTCTCCATCTCCAGTTCTGGACAGTGCAG GTTGCCAAGAAAAGGGACCTAGTTTAGCCATTGCCCTGGAGATGAAATTAATGGAGGCTCAA GGATAGATGAGCTCTGAGTTTCTCAGTACTCCCTCAAGACTGGACATCTTGGTCTTTTTCTC GAGGTGGGGGGGGGGGGGGTATATTGGAACTCTTCTAACCTCCTTGGGCTATATTTTCTC  ${\tt TCCTCGAGTTGCTCCTCATGGCTGGGCTCATTTCGGTCCCTTTCTCCTTGGTCCCAGACCTT}$ GGGGGAAAGGAAGTGCATGTTTGGGAACTGGCATTACTGGAACTAATGGTTTTAACCT CCTTAACCACCAGCATCCCTCCTCCCCAAGGTGAAGTGGAGGGTGCTGTGGTGAGCTGGC CACTCCAGAGCTGCAGTGCCACTGGAGGAGTCAGACTACCATGACATCGTAGGGAAGGAGGG GAGATTTTTTTGTAGTTTTTAATTGGGGTGTGGGAGGGCGGGGGAGGTTTTCTATAAACTGT ATCATTTTCTGCTGAGGGTGGAGTGTCCCATCCTTTTAATCAAGGTGATTGTGATTTTGACT AAAAAAAAAAAAAAAAAAAAAAA

# FIGURE 220

MGAAVFFGCTFVAFGPAFALFLITVAGDPLRVIILVAGAFFWLVSLLLASVVWFILVHVTDR SDARLQYGLLIFGAAVSVLLQEVFRFAYYKLLKKADEGLASLSEDGRSPISIRQMAYVSGLS FGIISGVFSVINILADALGPGVVGIHGDSPYYFLTSAFLTAAIILLHTFWGVVFFDACERRR YWALGLVVGSHLLTSGLTFLNPWYEASLLPIYAVTVSMGLWAFITAGGSLRSIQRSLLCKD

# FIGURE 221

# FIGURE 222

GACCGACCGTTCAGATGCCCGGTTCCAGTACGGCTTCCTGATTTTTGGTGCTGCTGTNTCTG
TCCTTCTACAGGAGGTGTTCCGCTTTGCCTANTACAAGCTGCTTAAGAAGGCAGATGAGGGG
TTAGCATNGCTGAGTGAGGACGGAAGATCACCCATTTCCATCCGCCAGATGGCCTATGTTTN
TGGTNTTTCCTTCGGTATCATCAGTGGTGTTTTNTCTGTTATCAATATTTTGGNTGATGCAN
TTGGGCCAGGTGTGGTTGGGATCCATGGAGANTCACCCTATTAATTCCTGAATTCAGCCTTT
NTGACAGCAGCCATTATCCTGNTCCATACCTTTTGGGGAGTTGTGTTTTTTGATGCCTGTGA
GAGGAG

### FIGURE 223

NGTTGGAGAAGTGGCGCGGACNTTCATTTGGGGTTTCCCCCCTTTCCCTTTCCCCG
GGGTCTGGGGTGACATTGCACGGGCCCCTCGTGGGGTCGCGTTGCCACCCCACGCGGACTCC
CCAGNTGGNGCGCCCTTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCCCTTCCCACNTG
ACCAGCCATGGGGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTCGCGTTCGGCCCGGCCTTCG
CGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGCA
TTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCATCTTTGGTCCATGTGAC
CGACCGGTCAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTCC
TTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTTA
GCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGG
TCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTTG
GGCCAGGTGTGGTTGGGATCCATGGAGACTCACCC

### FIGURE 224

GTAAAAGAAAGTGGCCGGACCTTCATTGGGGTTTCGGTTCCCCCCTTTCCCNTTCCCCGGGG
TCTGGGGGTGACATTGCACCGCGCCCNTCGTGGGGTCGCGTTGCCACCCCACGCGGACTCCC
CAGNTGGCGCGCCCCCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCCCTTCCCACCTGA
CCAGCCATGGGGGGCTGCGGTGTTTTTCGGGCTGCACTTTCGTCGCGTTCGGGCCCGGCCTTC
GCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGC
ATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCATCTTTGGTCCATGTGA
CCGACCGGTCAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTC
CTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTT
AGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTG
GTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTT
GGGCCAGGTGTGGTTGGGATCCATGGAGAC

### FIGURE 225

GCCCCAGGGAGCAGTGGTTATAACTCAGGCCCGGTGCCCAGAGCCCAGGAGGAGGAGGCAG TGGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCCACCCC TGTCTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCCTCCTTGCCCTGTCTGGAGGCTGCT AGACTCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCGT CCTTGTGGTTCCTCTACCTGGGGAAATAAGGTGCAGCGGCC ATG GCTACAGCAAGACCCCCCTGGATGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGGTCACAGAGCATGTT CTCGCCAACAATGATGTTTCCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCA GGACCTGGGAGCTGGGCCGGGGAAGACGCCCGGTCGGATGACAGCAGCAGCCGCATCATCA ATGGATCCGACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAAC CAGCTCTACTGCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCAG GAAGAAGTTTTCAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGC AGCAGATGTTCCAGGGGGTCAAATCCATCCCCCACCCTGGCTACTCCCACCCTGGCCACTCT AACGACCTCATGCTCAAACTGAACAGAAGAATTCGTCCCACTAAAGATGTCAGACCCAT  $\verb|CCAAGAGCCCCCAAGTGCACTTCCCTAAGGTCCTCCAGTGCTTGAATATCAGCGTGCTAAGT|\\$ CAGAAAAGGTGCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGA CAAAGCAGGTAGAGACTCCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCC TGCAGGGACTCGTGTCCTGGGGAGATTACCCTTGTGCCCGGCCCAACAGACCGGGTGTCTAC ACGAACCTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC**TGA**GTCAT CCCAGGACTCAGCACCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTCAG ACCCTCATTCCTTCCCAGAGATGTTGAGAATGTTCATCTCTCCAGCCCCTGACCCCATGTCT CCTGGACTCAGGGTCTGCTTCCCCCACATTGGGCTGACCGTGTCTCTAGTTGAACCCTGG GAACAATTTCCAAAACTGTCCAGGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCAT CCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAA CTGAGAAGTGGAAAAAAAA

# FIGURE 226

MATARPPWMWVLCALITALLLGVTEHVLANNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDD SSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLS PVYESGQQMFQGVKSIPHPGYSHPGHSNDLMLIKLNRRIRPTKDVRPINVSSHCPSAGTKCL VSGWGTTKSPQVHFPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAGDKAGRDSCQGDSGGP VVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

### FIGURE 227

**ATG**GTCAACGACCGGTGGAAGACCATGGGCGGCGCTGCCCAACTTGAGGACCGGCCGCGA <u>CAAGCCGCAGCGGCCGAGCTGCGGCTACGTGCTGTGCACCGTGCTGCTGGCCCTGGCTGTGC</u> CAGCCACGGCTGGTGGGCGACCAGGAGCAGGAGCTGCTGGACACGCTGGCCGACCAGCTGCC CCGGCTGCTGGCCCGAGCCTCAGAGCTGCAGACGGAGTGCATGGGGCTGCGGAAGGGGCATG CCAGGTGTACTGTGACATGCGCACGGACGGCGGCGGCTGGACGGTGTTTCAGCGCCGGGAGG ACGCTCCGTGAACTTCTTCCGGGGCTGGGACGCGTACCGAGACGGCTTTGGCAGGCTCACC GGGGAGCACTGGCTAGGGCTCAAGAGGATCCACGCCCTGACCACACAGGCTGCCTACGAGCT GCACGTGGACCTGGAGGACTTTGAGAATGCACGGCCTATGCCCGCTACGGGAGCTTCGGCG TGGGCTTGTTCTCCGTGGACCCTGAGGAAGACGGGTACCCGCTCACCGTGGCTGACTATTCC GGCACTGCAGGCGACTCCCTCCTGAAGCACAGCGGCATGAGGTTCACCACCAAGGACCGTGA CAGCGACCATTCAGAGAACAACTGTGCCGCCTTCTACCGCGGTGCCTGGTGGTACCGCAACT GGACTGCGCTAATGGGAAGCTCTTGGTTTTCTGGGCTGGGGCCTAGGCAGGGCTGGGATGAG GGTGCCTAGGGGGTGTGGGGTTCCGTTCTCCCTTCCCCTCCACTGAAGTTTGTGCTTAAAA AACAATAAATTTGACTTGGCACCACTGGGGGTTGGTGGGAGAGGCCGTGTGACCTGGCTCTC TGTCCCAGTGCCACCAGGTCATCCACATGCGCAG

# FIGURE 228

MVNDRWKTMGGAAQLEDRPRDKPQRPSCGYVLCTVLLALAVLLAVAVTGAVLFLNHAHAPGT
APPPVVSTGAASANSALVTVERADSSHLSILIDPRCPDLTDSFARLESAQASVLQALTEHQA
QPRLVGDQEQELLDTLADQLPRLLARASELQTECMGLRKGHGTLGQGLSALQSEQGRLIQLL
SESQGHMAHLVNSVSDILDALQRDRGLGRPRNKADLQRAPARGTRPRGCATGSRPRDCLDVL
LSGQQDDGVYSVFPTHYPAGFQVYCDMRTDGGGWTVFQRREDGSVNFFRGWDAYRDGFGRLT
GEHWLGLKRIHALTTQAAYELHVDLEDFENGTAYARYGSFGVGLFSVDPEEDGYPLTVADYS
GTAGDSLLKHSGMRFTTKDRDSDHSENNCAAFYRGAWWYRNCHTSNLNGQYLRGAHASYADG
VEWSSWTGWQYSLKFSEMKIRPVREDR

### FIGURE 229

GCAGTCAGAGACTTCCCCTGCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCT TGCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCCAGGGCAATCCGACCACATTTCACTCT CACCGCTGTAGGAATCCAG**ATG**CAGGCCAAGTACAGCACGAGGGACATGCTGGATGATG ATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACTCGGCATCCAGAGCCC CGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGAC TTTGTGCTTGGTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTCAGTACTACC CAAGAGTTGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGC TGAAAAACTCTGTCGTGAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGTACAG AACAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAG CCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTT TGCGCCCTGACAGTGGCAAGGCCTGGCTGTGGATGGAACCCCTTTCACTTCTGAACTG TTCCATATTATAATAGATGTCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGG GATGATCTTCTCAAAGGACTGCAAAGAATTGAAGCGTTGTGTCTGTGAGAGAAGGGCAGGAA TGGTGAAGCCAGAGAGCCTCCATGTCCCCCCTGAAACATTAGGCGAAGGTGAC**TGA**TTCGCC CTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACAT TGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCTCGTG TTTCCTGTTCAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACA AGAAGTCTTATTTACATGCCACCAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTG GCTTAGAGATAACTTTTAGCTCTCTTTTCTTCTCAATGTCTAATATCACCTCCCTGTTTTCAT GTCTTCCTTACACTTGGTGGAATAAGAAACTTTTTGAAGTAGAGGAAATACATTGAGGTAAC ATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTACC AGCAAATACACAAGGAATTCTTTTTGTTTGTTTCAGTTCATACTAGTCCCTTCCCAATCCAT CAGTAAAGACCCCATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAG AATCTCAAATCTCAATGCCTTATAAGCATTCCTTCCTGTGTCCATTAAGACTCTGATAATTG TCTCCCCTCCATAGGAATTTCTCCCAGGAAAGAAATATATCCCCATCTCCGTTTCATATCAG AACTACCGTCCCCGATATTCCCTTCAGAGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCA TCTGCACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTCAG GTACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

# FIGURE 230

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCLVLL
IGLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRE
LYNKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAAS
QSYSEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMIFSKD
CKELKRCVCERRAGMVKPESLHVPPETLGEGD

### FIGURE 231

### FIGURE 232

GCCGAGCGCAAGAACCCTGCGCAGCCCAGAGCAGCTGCTGGAGGGGAATCGAGGCGCGCTC CCCGAGCCCTCCGGATCCGCCCCCTCCCGGTCCCGCCCCCTCGGAGACTCCTCTGGCTGCT TCGGTGCTGCGGCCCGCAGGGCCCGTGGCCGTGGGCATCTCCCCTGGGCCTTCACCCTGAGCCT GCTCAGCGTCACCTGGGTGGAGGAGCCGTGCGGCCCAGGCCCCCAACCTGGAGACTCTG AGCTGCCGCCGCGCGCAACACCAACGCGGCGCGCCCCAACTCGGTGCAGCCCGGAGCG GAGCGCGAGAAGCCCGGGGCCGAAGGCGCCGGGGAGAATTGGGAGCCGCGCGTCTTGCC CTACCACCCTGCACAGCCCGGCCAGGCCGCCAAAAAGGCCGTCAGGACCCGCTACATCAGCA CGGAGCTGGGCATCAGGCAGAGGCTGCTGGTGGCGGTGCTGACCTCTCAGACCACGCTGCCC ACGCTGGGCGTGGCCGTGAACCGCACGCTGGGGCACCGGCTGGAGCGTGTGGTGTTCCTGAC GGGCGCACGGGGCCCCGGGCCCCACCTGGCATGGCAGTGGTGACGCTGGGCGAGGAGCGAC CCATTGGACACCTGCACCTGCGCGCCCACCTGCTGGAGCACGACGACGACTTTGAC TGGTTCTTCCTGGTGCCTGACACCACCTACACCGAGGCGCACGGCCTGGCACGCCTAACTGG GAGAGCCCACCCCGGCCGCTACTGCCACGGAGGCTTTGGGGGTGCTGCTGTCGCGCATGCTG CTGCAACAACTGCGCCCCCACCTGGAAGGCTGCCGCAACGACATCGTCAGTGCGCGCCCTGA CGAGTGGCTGGGTCGCTGCATTCTCGATGCCACCGGGGTGGGCTGCACTGGTGACCACGAGG GGGTGCACTATAGCCATCTGGAGCTGAGCCCTGGGGAGCCAGTGCAGGAGGGGGACCCTCAT TTCCGAAGTGCCCTGACAGCCCACCCTGTGCGTGACCCTGTGCACATGTACCAGCTGCACAA AGCTTTCGCCCGAGCTGAACTGGAACGCACGTACCAGGAGATCCAGGAGTTACAGTGGGAGA TCCAGAATACCAGCCATCTGGCCGTTGATGGGGACCGGGCAGCTGCTTGGCCCGTGGGTATT CCAGCACCATCCCGCCCGGCCTCCCGCTTTGAGGTGCTGCGCTGGGACTACTTCACGGAGCA GCACGCTTTCTCCTGCGCCGATGGCTCACCCCGCTGCCCACTGCGTGGGGCTGACCGGGCTG ATGTGGCCGATGTTCTGGGGACAGCTCTAGAGGAGCTGAACCGCCGCTACCACCCGGCCTTG  $\tt CGGCTCCAGAAGCAGCTGGTGAATGGCTACCGACGCTTTGATCCGGCCCGGGGTATGGA$ ATACACGCTGGACTTGCAGCTGGAGGCACTGACCCCCCAGGGAGGCCGCCGGCCCCTCACTC GCCGAGTGCAGCTGCTCCGGCCGCTGAGCCGCGTGGAGATCTTGCCTGTGCCCTATGTCACT GAGGCCTCACGTCTCACTGTGCTGCCTCTAGCTGCGGCTGAGCGTGACCTGGCCCCTGG CTTCTTGGAGGCCTTTGCCACTGCAGCACTGGAGCCTGGTGATGCTGCGGCAGCCCTGACCC TGCTGCTACTGTATGAGCCGCGCCAGGCCCAGCGCGTGGCCCATGCAGATGTCTTCGCACCT GTCAAGGCCCACGTGGCAGAGCTGGAGCGGCGTTTCCCCGGTGCCCGGGTGCCATGGCTCAG TGTGCAGACAGCCGCACCCTCACCACTGCGCCTCATGGATCTACTCTCCAAGAAGCACCCGC TGGACACACTGTTCCTGCTGGCCGGGCCAGACACGGTGCTCACGCCTGACTTCCTGAACCGC TGCCGCATGCATGCCATCTCCGGCTGGCAGGCCTTCTTTCCCATGCATTTCCAAGCCTTCCA CCCAGGTGTGGCCCCACCACAAGGGCCTGGGCCCCCAGAGCTGGGCCGTGACACTGGCCGCT TTGATCGCCAGGCAGCCAGCGAGGCCTGCTTCTACAACTCCGACTACGTGGCAGCCCGTGGG CGCCTGGCGGCAGCCTCAGAACAAGAAGAGGAGCTGCTGGAGAGCCTGGATGTACGAGCT GTTCCTCCACTTCTCCAGTCTGCATGTGCTGCGGGGCGGTGGAGCCGGCGCTGCTGCAGCGCT ACCGGGCCCAGACGTGCAGCGCGAGGCTCAGTGAGGACCTGTACCACCGCTGCCTCCAGAGC GTGCTTGAGGGCCTCGGCTCCCGAACCCAGCTGGCCATGCTACTCTTTGAACAGGAGCAGGG  ${\tt CAACAGCACC} \underline{{\tt TGA}} {\tt CCCCACCCTGTCCCCGTGGGCCGTGGCATGGCCACACCCCACCTT}$  $\mathtt{CTCCCCAAAACC}$  AGAGCCACCTGCCAGCCTCGCTGGGCAGGGCTGGCCGTAGCCAGACCCC AAGCTGGCCCACTGGTCCCCTCTCTGGCTCTGTGGGTCCCTGGGCTCTGGACAAGCACTGGG GGACGTGCCCCAGAGCCACCCACTTCTCATCCCAAACCCAGTTTCCCTGCCCCCTGACGCT GCTGATTCGGGCTGTGGCCTCCACGTATTTATGCAGTACAGTCTGCCTGACGCCAGCCCTGC CTCTGGGCCCTGGGGCTGTAGAAGAGTTGTTGGGGAAGGAGGAGCTGAGGAGGGG GCATCTCCCAACTTCTCCCTTTTGGACCCTGCCGAAGCTCCCTGCCTTTAATAAACTGGCCA AGTGTGGAAAAA

### FIGURE 233

MRASLLLSVLRPAGPVAVGISLGFTLSLLSVTWVEEPCGPGPPQPGDSELPPRGNTNAARRP
NSVQPGAEREKPGAGEGAGENWEPRVLPYHPAQPGQAAKKAVRTRYISTELGIRQRLLVAVL
TSQTTLPTLGVAVNRTLGHRLERVVFLTGARGRRAPPGMAVVTLGEERPIGHLHLALRHLLE
QHGDDFDWFFLVPDTTYTEAHGLARLTGHLSLASAAHLYLGRPQDFIGGEPTPGRYCHGGFG
VLLSRMLLQQLRPHLEGCRNDIVSARPDEWLGRCILDATGVGCTGDHEGVHYSHLELSPGEP
VQEGDPHFRSALTAHPVRDPVHMYQLHKAFARAELERTYQEIQELQWEIQNTSHLAVDGDRA
AAWPVGIPAPSRPASRFEVLRWDYFTEQHAFSCADGSPRCPLRGADRADVADVLGTALEELN
RRYHPALRLQKQQLVNGYRRFDPARGMEYTLDLQLEALTPQGGRRPLTRRVQLLRPLSRVEI
LPVPYVTEASRLTVLLPLAAAERDLAPGFLEAFATAALEPGDAAAALTLLLLYEPRQAQRVA
HADVFAPVKAHVAELERRFPGARVPWLSVQTAAPSPLRLMDLLSKKHPLDTLFLLAGPDTVL
TPDFLNRCRMHAISGWQAFFPMHFQAFHPGVAPPQGPGPPELGRDTGRFDRQAASEACFYNS
DYVAARGRLAAASEQEEELLESLDVYELFLHFSSLHVLRAVEPALLQRYRAQTCSARLSEDL
YHRCLQSVLEGLGSRTQLAMLLFEQEQGNST

### FIGURE 234

GCTCTGGCCGGCCCGGCGATTGGTCACCGCCCGCTAGGGGACAGCCCTGGCCTCCTCTGAT TGGCAAGCGCTGGCCACCTCCCCACACCCCTTGCGAACGCTCCCCTAGTGGAGAAAAGGAGT AGCTATTAGCCAATTCGGCAGGGCCCGCTTTTTAGAAGCTTGATTTCCTTTGAAGATGAAAG ACTAGCGGAAGCTCTGCCTCTTTCCCCAGTGGGCGAGGGAACTCGGGGCGATTGGCTGGGAA CTGTATCCACCCAAATGTCACCGATTTCTTCCTATGCAGGAAATGAGCAGACCCATCAATAA GAAATTTCTCAGCCTGGCCGAAAATGGTTGGCCCCACGAAGCCACGACAACTGGAGGCAAAG AGGGTTGCTCAACGCCCCGCCTCATTGGAAAACCAAATCAGATCTGGGACCTATATAGCGTG GCGGAGGCGGGCGATGATTGTCGCGCTCGCACCCACTGCAGCTGCGCACAGTCGCATTTCT ATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCTGTGGCGGGAAAGCGGCCCCCAGAAC CGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGACGAAAACTTGTATGAGAAGAACCCA GACTCCCATGGTTATGACAAGGACCCCGTTTTGGACGTCTGGAACATGCGACTTGTCTTCTT CTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACCTTTGTGGCCTATCTGCCTGACTACA GGATGAAAGAGTGGTCCCGCCGCAAGCTGAGAGGCTTGTGAAATACCGAGAGGCCAATGGC CTTCCCATCATGGAATCCAACTGCTTCGACCCCAGCAAGATCCAGCTGCCAGAGGATGAGTG CTCTTCTCAGAGCACCTAATTAAAGGGGCTGAAAGTCTGAA

# FIGURE 235

MAAGLFGLSARRLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPPEPTTPWQEDPEPE DENLYEKNPDSHGYDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAER LVKYREANGLPIMESNCFDPSKIQLPEDE

## FIGURE 236

### FIGURE 237

TGCAGAACCCCCACGCGACAGCCTGCGGGAGGAACTTGTCATCACCCCGCTGCCTTCCGGGG ACGTAGCCGCCACATTCCAGTTCCGCACGCGCTGGGATTCGGAGCTTCAGCGGGAAGGAGTG TCCCATTACAGGCTCTTTCCCAAAGCCCTGGGGCAGCTGATCTCCAAGTATTCTCTACGGGA GCTGCACCTGTCATTCACACAAGGCTTTTGGAGGACCCGATACTGGGGGCCCACCCTTCCTGC AGGCCCCATCAGGTGCAGAGCTGTGGGTCTGGTTCCAAGACACTGTCACTGATGTGGATAAA TCTTGGAAGGAGCTCAGTAATGTCCTCTCAGGGATCTTCTGCGCCCTCTCTCAACTTCATCGA CTCCACCAACACAGTCACTCCCACTGCCTCCTTCAAACCCCTGGGTCTGGCCAATGACACTG ACCACTACTTTCTGCGCTATGCTGTGCTGCCGCGGGAGGTGGTCTGCACCGAAAACCTCACC CCCTGGAAGAAGCTCTTGCCCTGTAGTTCCAAGGCAGGCCTCTCTGTGCTGCTGAAGGCAGA TCGCTTGTTCCACACCAGCTACCACTCCCAGGCAGTGCATATCCGCCCTGTTTGCAGAAATG CACGCTGTACTAGCATCTCCTGGGAGCTGAGGCAGACCCTGTCAGTTGTATTTGATGCCTTC ATCACGGGGCAGGGAAAGAAGACTGGTCCCTCTTCCGGATGTTCTCCCGAACCCTCACGGA GCCCTGCCCCTGGCTTCAGAGAGCCGAGTCTATGTGGACATCACCACCTACAACCAGGACA ACGAGACATTAGAGGTGCACCCACCCCGACCACTACATATCAGGACGTCATCCTAGGCACT CGGAAGACCTATGCCATCTATGACTTGCTTGACACCGCCATGATCAACAACTCTCGAAAACCT CAACATCCAGCTCAAGTGGAAGAGACCCCCAGAGAATGAGGCCCCCCAGTGCCCTTCCTGC ATGCCCAGCGGTACGTGAGTGGCTATGGGCTGCAGAAGGGGGAGCTGAGCACACTGCTGTAC AACACCCACCCATACCGGGCCTTCCCGGTGCTGCTGCTGGACACCGTACCCTGGTATCTGCG GCTGTATGTGCACACCCTCACCATCACCTCCAAGGGCAAGGAGAACAAACCAAGTTACATCC ACTACCAGCCTGCCAGGACCGGCTGCAACCCCACCTCCTGGAGATGCTGATTCAGCTGCCG GCCAACTCAGTCACCAAGGTTTCCATCCAGTTTGAGCGGGCGCTGCTGAAGTGGACCGAGTA CACGCCAGATCCTAACCATGGCTTCTATGTCAGCCCATCTGTCCTCAGCGCCCTTGTGCCCA GCATGGTAGCAGCCAAGCCAGTGGACTGGGAAGAGAGTCCCCTCTTCAACAGCCTGTTCCCA GTCTCTGATGGCTCTAACTACTTTGTGCGGCTCTACACGGAGCCGCTGCTGGTGAACCTGCC GACACCGGACTTCAGCATGCCCTACAACGTGATCTGCCTCACGTGCACTGTGGTGGCCGTGT GCTACGGCTCCTTCTACAATCTCCTCACCCGAACCTTCCACATCGAGGAGCCCCGCACAGGT GGCCTGGCCAAGCGGCTGGCCAACCTTATCCGGCGCGCCCGAGGTGTCCCCCCACTC**TGA**TT CTTGCCCTTTCCAGCAGCTGCAGCTGCCGTTTCTCTCTGGGGAGGGGAGCCCAAGGGCTGTT TCTGCCACTTGCTCTCAGAGTTGGCTTTTGAACCAAAGTGCCCTGGACCAGGTCAGGGC CTACAGCTGTGTTGTCCAGTACAGGAGCCACGAGCCAAATGTGGCATTTGAATTTAA CTTAGAAATTCATTTCCTCACCTGTAGTGGCCACCTCTATATTGAGGTGCTCAATAAGCAAA AGTGGTCGGTGGCTGTATTGGACAGCACAGAAAAAGATTTCCATCACCACAGAAAGGTC GGCTGGCAGCACTGGCCAAGGTGATGGGGTGTGCTACACAGTGTATGTCACTGTAGTGGA 

### FIGURE 238

MPLALLVLLLLGPGGWCLAEPPRDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHY
RLFPKALGQLISKYSLRELHLSFTQGFWRTRYWGPPFLQAPSGAELWVWFQDTVTDVDKSWK
ELSNVLSGIFCASLNFIDSTNTVTPTASFKPLGLANDTDHYFLRYAVLPREVVCTENLTPWK
KLLPCSSKAGLSVLLKADRLFHTSYHSQAVHIRPVCRNARCTSISWELRQTLSVVFDAFITG
QGKKDWSLFRMFSRTLTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTYQDVILGTRKT
YAIYDLLDTAMINNSRNLNIQLKWKRPPENEAPPVPFLHAQRYVSGYGLQKGELSTLLYNTH
PYRAFPVLLLDTVPWYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANS
VTKVSIQFERALLKWTEYTPDPNHGFYVSPSVLSALVPSMVAAKPVDWEESPLFNSLFPVSD
GSNYFVRLYTEPLLVNLPTPDFSMPYNVICLTCTVVAVCYGSFYNLLTRTFHIEEPRTGGLA
KRLANLIRRARGVPPL

### FIGURE 239

# FIGURE 240

 ${\tt MGSSSFLVLMVSLVLVTLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC} \\ {\tt YLHCGFKCVIPVKELEEGGNKDEDVSRPYPEPGWEAKCPGSSSTRCPQK} \\$ 

### Signal sequence:

amino acids 1-19

### N-myristoylation sites:

amino acids 23-29, 27-33, 32-38, 102-108

### WAP-type 'four-disulfide core' domain signature:

amino acids 49-63

### FIGURE 241

AAACTCAGCACTTGCCGGAGTGGCTCATTGTTAAGACAAAGGGTGTGCACTTCCTGGCCAGG AAACCTGAGCGGTGAGACTCCCAGCTGCCTACATCAAGGCCCCAGGACATGCAGAACCTTCC TCTAGAACCCGACCCACCACCATGAGGTCCTGCCTGTGGAGATGCAGGCACCTGAGCCAAGG CGTCCAGTGGTCCTTGCTTCTGGCTGTCCTGGTCTTCTTCTCTTCGCCTTGCCCTCTTTTA CTACAGTCCCTGGCAAAGCCTAAGTCCCAGGCACCCACAAGGGCGAGGAGGACAACCATCTA TGCAGAGCCAGCGCCAGAGAACAATGCCCTCAACACACAAACCCAGCCCAAGGCCCACACCA ACAGCACAGAGGGCAGCATGGAAGAGCCCAGAAAAAGAGAAAACCATGGTGAACACACTGTC ACCCAGAGGGCAAGATGCAGGGATGGCCTCTGGCAGGACAGAGGCACAATCATGGAAGAGCC AGGACACAAAGACGACCCAAGGAAATGGGGGCCAGACCAGGAAGCTGACGGCCTCCAGGACG GTGTCAGAGAAGCACCAGGGCAAAGCGGCAACCACAGCCAAGACGCTCATTCCCAAAAGTCA GCACAGAATGCTGGCTCCCACAGGAGCAGTGTCAACAAGGACGAGACAGAAAGGAGTGACCA CAGCAGTCATCCCACCTAAGGAGAAGAAACCTCAGGCCACCCCACCCCCTGCCCCTTTCCAG AGCCCCACGACGCAGAGAACCAAAGACTGAAGGCCGCCAACTTCAAATCTGAGCCTCGGTG GGATTTTGAGGAAAAATACAGCTTCGAAATAGGAGGCCTTCAGACGACTTGCCCTGACTCTG TGAAGATCAAAGCCTCCAAGTCGCTGTGGCTCCAGAAACTCTTTCTGCCCAACCTCACTCTC TTCCTGGACTCCAGACACTTCAACCAGAGTGAGTGGGACCGCCTGGAACACTTTGCACCACC CTTTGGCTTCATGGAGCTCAACTACTCCTTGGTGCAGAAGGTCGTGACACGCTTCCCTCCAG TGCCCCAGCAGCAGCTGCTCCTGGCCAGCCTCCCCGCTGGGAGCCTCCGGTGCATCACCTGT GCCGTGGTGGCCAACGGGGCCATCCTGAACAACTCCCACATGGGCCAGGAGATAGACAGTCA CGACTACGTGTTCCGATTGAGCGGAGCTCTCATTAAAGGCTACGAACAGGATGTGGGGACTC GGACATCCTTCTACGGCTTTACCGCCTTCTCCCTGACCCAGTCACTCCTTATATTGGGCAAT CCGGGACTATGAGTGGCTGGAAGCACTGCTTATGAATCAGACGGTGATGTCAAAAAACCTTT TCTGGTTCAGGCACAGACCCCAGGAAGCTTTTCGGGAAGCCCTGCACATGGACAGGTACCTG TTGCTGCACCCAGACTTTCTCCGATACATGAAGAACAGGTTTCTGAGGTCTAAGACCCTGGA TGGTGCCCACTGGAGGATATACCGCCCCACCACTGGGGCCCTCCTGCTGCTCACTGCCCTTC AGCTCTGTGACCAGGTGAGTGCTTATGGCTTCATCACTGAGGGCCCATGAGCGCTTTTCTGAT CACTACTATGATACATCATGGAAGCGGCTGATCTTTTACATAAACCATGACTTCAAGCTGGA GAGAGAAGTCTGGAAGCGGCTACACGATGAAGGGATAATCCGGCTGTACCAGCGTCCTGGTC  ${\tt CCGGAACTGCCAAAGCCAAGAAC{\tt TGA}}{\tt CCGGGGGCCAGGGCTGCCATGGTCTCCTTGCCTGCTC}$ CAAGGCACAGGATACAGTGGGAATCTTGAGACTCTTTGGCCATTTCCCATGGCTCAGACTAA GCTCCAAGCCCTTCAGGAGTTCCAAGGGAACACTTGAACCATGGACAAGACTCTCTCAAGAT GGCAAATGGCTAATTGAGGTTCTGAAGTTCTTCAGTACATTGCTGTAGGTCCTGAGGCCAGG GATTTTTAATTAAATGGGGTGATGGGTGGCCAATACCACAATTCCTGCTGAAAAAACACTCTT CCAGTCCAAAAGCTTCTTGATACAGAAAAAAGAGCCTGGATTTACAGAAACATATAGATCTG GTTTGAATTCCAGATCGAGTTTACAGTTGTGAAATCTTGAAGGTATTACTTAACTTCACTAC AGATTGTCTAGAAGACCTTTCTAGGAGTTATCTGATTCTAGAAGGGTCTATACTTGTCCTTG TCTTTAAGCTATTTGACAACTCTACGTGTTGTAGAAAACTGATAATAATACAAATGATTGTT 

### FIGURE 242

MRSCLWRCRHLSQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKP
KSQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAW
KSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNGGQTRKLTASRTVSEKHQG
KAATTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQSPTTQRN
QRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDSRHF
NQSEWDRLEHFAPPFGFMELNYSLVQKVVTRFPPVPQQQLLLASLPAGSLRCITCAVVGNGG
ILNNSHMGQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGFTAFSLTQSLLILGNRGFKNVP
LGKDVRYLHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFL
RYMKNRFLRSKTLDGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYYDTSW
KRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

### Cytoplasmic Domain:

amino acids 1-10

#### Type II Transmembrane Domain:

amino acids 11-35

#### Lumenal catalytic Domain:

amino acids 36-600

### Ribonucleotide Reductase small subunit Signature:

amino acids 481-496

#### N-glycosylation Sites:

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

## FIGURE 243

# FIGURE 244

 ${\tt MRGPGHPLLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKHFHRLYHNT} \\ {\tt RDCTIPAYYKRCARLLTRLAVSPVCMEDK}$ 

## FIGURE 245

GGGCTGGGCCCGCCGCAGCTCCAGCTGGCCGGCTTGGTCCTGCGGTCCCTTCTCTGGGAGG CTCACCGCAATCGCTCTGTTGGTGCTGGGGGGCTCCCCTGGTGCTGGCCGGCGAGGACTGCCT GTGGTACCTGGACCGGAATGGCTCCTGGCATCCGGGGTTTAACTGCGAGTTCTTCACCTTCT GCTGCGGGACCTGCTACCATCGGTACTGCTGCAGGGACCTGACCTTGCTTATCACCGAGAGG CAGCAGAAGCACTGCCTGGCCTTCAGCCCCAAGACCATAGCAGGCATCGCCTCAGCTGTGAT CCTCTTTGTTGCTGTTGCCACCACCATCTGCTGCTTCCTCTGTTCCTGTTGCTACCTGT ACCGCCGGCGCCAGCAGCTCCAGAGCCCATTTGAAGGCCAGGAGATTCCAATGACAGGCATC CCAGTGCAGCCAGTATACCCATACCCCAGGACCCCAAAGCTGGCCCTGCACCCCCACAGCC TGGCTTCATGTACCCACCTAGTGGTCCTGCTCCCCAATATCCACTCTACCCAGCTGGGCCCC CAGTCTACAACCCTGCAGCTCCTCCTCCTATATGCCACCACAGCCCTCTTACCCGGGAGCC TGAGGAACCAGCCATGTCTCTGCTGCCCCTTCAGTGATGCCAACCTTGGGAGATGCCCTCAT CCTGTACCTGCATCTGGTCCTGGGGGTGGCAGGAGTCCTCCAGCCACCAGGCCCCAGACCAA GCCAAGCCCTGGGCCCTACTGGGGACAGAGCCCCAGGGAAGTGGAACAGGAGCTGAACTAGA ACTATGAGGGGTTGGGGGGGGGCTTGGAATTATGGGCTATTTTTACTGGGGGCAAGGGAGG GAGATGACAGCCTGGGTCACAGTGCCTGTTTTCAAATAGTCCCTCTGCTCCCAAGATCCCAG CCAGGAAGGCTGGGGCCCTACTGTTTGTCCCCTCTGGGCTGGGGTGGGGGGAGGAGGAGGT TCCGTCAGCAGCTGGCAGTAGCCCTCCTCTCTGGCTGCCCCACTGGCCACATCTCTGGCCTG CTAGATTAAAGCTGTAAAGACAAAA

# FIGURE 246

MPPAGLRRAAPLTAIALLVLGAPLVLAGEDCLWYLDRNGSWHPGFNCEFFTFCCGTCYHRYC CRDLTLLITERQQKHCLAFSPKTIAGIASAVILFVAVVATTICCFLCSCCYLYRRRQQLQSP FEGQEIPMTGIPVQPVYPYPQDPKAGPAPPQPGFMYPPSGPAPQYPLYPAGPPVYNPAAPPP YMPPQPSYPGA

### Transmembrane Domains:

amino acids 10-28, 85-110

### N-glycosylation Site:

amino acids 38-41

### N-myristoylation Sites:

amino acids 5-10, 88-93

### FIGURE 247

GGGGGAGCTAGGCCGGCGGCAGTGGTGGTGGCGCCGCGCAAGGGTGAGGGCGCCCCAGAA  $\tt CCCCAGGTAGGTAGAGCAAGAAGAAGGTGTTTCTGCCCCTCAAATGGTCCCTTGCAACCATG$ TCATTTCTACTTTCCTCACTGTTGGCTCTCTTAACTGTGTCCACTCCTTGAACCATG
TCATTTCTACTTTCCTCACTGTTGGCTCTCTTAACTGTGTCCACTCCTTCATGGTGTCAGAG
CACTGAAGCATCTCCAAAACGTAGTGATGGGACACCATTTCCTTGGAATAAAATACGACTTC
CTGAGTACGTCATCCCAGTTCATTATGATCTCTTGATCCATGCAAACCTTACCACGCTGACC
TTCTGGGGAACCACGAAAGTAGAAATCACACACGCCACCACCACCACCATCATCCTGCA TAGTCACCACCTGCAGATATCTAGGGCCACCCTCAGGAAGGGAGCTGGAGAGAGGCTATCGG AAGAACCCTGCAGGTCCTGGAACACCCCCCTCAGGAGCAAATTGCACTGCTGGCTCCCGAGCCCCTCCTTGTCGGGCTCCCGTACACAGTTGTCATTCACTATGCTGGCAATCTTTCGGAGAC ATTGGTGAAATCTGTGACTGTTGCTGAAGGACTCATAGAAGACCATTTTGATGTCACTGTGA TAGGAGTATCTTAGGGCTGACGCATTTAAAAGIGGIAIIGIACAGIAICICCAGAAGCA
TAGCTATAAAAATACAAAAAACGAGGACCTGTGGGATAGTATGGCAAGTATTTGCCCTACAG
ATGGTGTAAAAGGGATGGATGGCTTTTGCTCTAGAAGTCAACATTCATCTTCATCCTCACAT
TGGCATCAGGAAGGGGTGGATGTGAAAACCATGATGAACACTTGGACACTGCAGAGGGGTTT
TCCCCTAATAACCATCACAGTGAGGGGGAGGAATGTACACATGAAGCAAGAGCACTACATGA
AGGGCTCTGACGGCGCCCCGGACACTGGGTACCTTGTGGCATTGATCACCC AGCAAATCCAACATGGTCCATCGATTTTTGCTAAAAACAAAAACAGATGTGCTCATCCTCCC TTCTCAGCTCCGTTGTGTCCAACAGACAATTGAAACCATTGAAGAAAACATCGGTTGGATGG ATAAGAATTTTGATAAAATCAGAGTGTGGCTGCAAAGTGAAAAGCTTGAACGTATG**TAA**AAA TTCCTCCCTTGCCCGGTTCCTGTTATCTCTAATCACCAACATTTTGTTGAGTGTATTTTCAA
ACTAGAGATGGCTGTTTTTGGCTCCAACTGGAGATACTTTTTTCCCTTCAACTCATTTTTTGA
CTATCCCTGTGAAAAGAATAGCTGTTAGTTTTTCATGAATGGGCTTTTTCATGAATGGGCTA
TCGCTACCATGTGTTTTGTTCATCACAGGTGTTGCCCTGCAACGTAAACCCAAGTGTTGGGT 

### FIGURE 248

MVFLPLKWSLATMSFLLSSLLALLTVSTPSWCQSTEASPKRSDGTPFPWNKIRLPEYVIPVH YDLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLQVLE HPPQEQIALLAPEPLLVGLPYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTQFEPTA ARMAFPCFDEPAFKASFSIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVA FIISDFESVSKITKSGVKVSVYAVPDKINQADYALDAAVTLLEFYEDYFSIPYPLPKQDLAA IPDFQSGAMENWGLTTYRESALLFDAEKSSASSKLGITVTVAHELAHQWFGNLVTMEWWNDL WLNEGFAKFMEFVSVSVTHPELKVGDYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFD DVSYDKGACILNMLREYLSADAFKSGIVQYLQKHSYKNTKNEDLWDSMASICPTDGVKGMDG FCSRSQHSSSSSHWHQEGVDVKTMMNTWTLQRGFPLITITVRGRNVHMKQEHYMKGSDGAPD TGYLWHVPLTFITSKSNMVHRFLLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSL TGLLKGTHTAVSSNDRASLINNAFQLVSIGKLSIEKALDLSLYLKHETEIMPVFQGLNELIP MYKLMEKRDMNEVETQFKAFLIRLLRDLIDKQTWTDEGSVSEQMLRSELLLLACVHNYQPCV QRAEGYFRKWKESNGNLSLPVDVTLAVFAVGAQSTEGWDFLYSKYQFSLSSTEKSQIEFALC RTQNKEKLQWLLDESFKGDKIKTQEFPQILTLIGRNPVGYPLAWQFLRKNWNKLVQKFELGS SSIAHMVMGTTNQFSTRTRLEEVKGFFSSLKENGSQLRCVQQTIETIEENIGWMDKNFDKIR VWLOSEKLERM

### Signal peptide:

amino acids 1-34

#### N-glycosylation sites:

amino acids 70-74, 154-158, 414-418, 760-764, 901-905

Neutral zinc metallopeptidases, zinc-binding region signature:

amino acids 350-360

### FIGURE 249

 $\texttt{CAGCCACAGACGGGTC} \underline{\textbf{ATG}} \texttt{AGCGCGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCCAC}$ TGCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCC GACCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGA CACGTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGG AGGCCAAGGACCAGGAGCCCCGCGTCACTGAGCACCGGATGGGCCCCCGGCCTCTCCCTGATC AAGGCTGTCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGAT CCAGCCAGGTTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCGTGGGTATGACTGAGA ACTGCAATAGGAAAGATTTTCTGACCTGTCATCGGGGGACCACCATTATGACACACGGAAAC TTGGCTCAAGAACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGT GTGTCAGGAGACGCTGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAG GCTGCAGCACTGTTGGGGCTCAAAATTCCCAGAAGACCACCATCCACTCAGCCCCTCCTGGG GTGCTTGTGGCCTCCTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAG CAGCGTTCTGCTGAACTCCCTCCTCAAGCTGCCCCTGTCCCAGGAGACCGGCAGTGTC CTACCTGTGTGCAGCCCCTTGGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGG GGCGCCACTCATTGTTATGATGGGTACATTCATCTCTCAGGAGGTGGGCTGTCCACCAAAAT GAGCATTCAGGGCTGCGTGGCCCAACCTTCCAGCTTCTTGTTGAACCACACCAGACAAATCG GGGGCTGAGGGCCTGGAGTCTCTCACTTGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTG  $\tt GTGGGGAGTGGTTTGCCCTTCCTGC\textbf{TAA} CTCTATTACCCCCACGATTCTTCACCGCTGCTGA$ CCACCCACACTCAACCTCCTCTGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTC ACACTGGGGAGAGCCTGGAGCATCCGGACTTGCCCTATGGGAGAGGGGACGCTGGAGGAGTG GCTGCATGTATCTGATAATACAGACCCTGTCCTTTCA

## FIGURE 250

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMLI ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP PADPGSLRCPVCLSMEGCLEGTTEEICPKGTTHCYDGLLRLRGGGIFSNLRVQGCMPQPGCN LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETL LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGC VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESLTWGVGLALAPALWWGVVC PSC

### FIGURE 251

 $\texttt{CAGG} \underline{\textbf{ATG}} \texttt{AGGGGGAATCTGGCCCTGGTGGGCGTTCTAATCAGCCTGGCCTTCCTGTCACTGCTG}$ CCATCTGGACATCCTCAGCCGGCTGGCGATGACGCCTGCTCTGTGCAGATCCTCGTCCCTGG CCTCAAAGGGGATGCGGGAGAGAGAGGGAGACAAAGGCGCCCCCGGACGCCTGGAAGAGTCG GCCCCACGGGAGAAAAGGACATGGGGGGACAAAGGACAGAAAGGCAGTGTGGGTCGTCAT GGAAAAATTGGTCCCATTGGCTCTAAAGGTGAGAAAGGAGATTCCGGTGACATAGGACCCCC AGATGGACAACCAGGTCTCTCAGCTGACCAGCGAGCTCAAGTTCATCAAGAATGCTGTCGCC GGTGTGCGCGAGACGGAGAGCAAGATCTACCTGCTGGTGAAGGAGGAGAAGCGCTACGCGGA CGCCCAGCTGTCCTGCCAGGGCCGCGGGGGCACGCTGAGCATGCCCAAGGACGAGGCTGCCA ATGGCCTGATGGCCGCATACCTGGCGCAAGCCGGCCTGGCCCGTGTCTTCATCGGCATCAAC GACCTGGAGAAGGAGGGCGCCTTCGTGTACTCTGACCACTCCCCCATGCGGACCTTCAACAA GTGGCGCAGCGGTGAGCCCAACAATGCCTACGACGAGGAGGACTGCGTGGAGATGGTGGCCT CGGGCGGCTGGAACGACGTGGCCTGCCACACCACCATGTACTTCATGTGTGAGTTTGACAAG GAGAACATG**TGA**GCCTCAGGCTGGGGCTGCCCATTGGGGGCCCCACATGTCCCTGCAGGGTT GGCAGGGACAGAGCCCAGACCATGGTGCCAGCCAGGGAGCTGTCCCTCTGTGAAGGGTGGAG GCTCACTGAGTAGAGGGCTGTTGTCTAAACTGAGAAAATGGCCTATGCTTAAGAGGAAAATG AAAGTGTTCCTGGGGTGCTGTCTCTGAAGAAGCAGAGTTTCATTACCTGTATTGTAGCCCCA ATGTCATTATGTAATTATTACCCAGAATTGCTCTTCCATAAAGCTTGTGCCTTTGTCCAAGC 

# FIGURE 252

MRGNLALVGVLISLAFLSLLPSGHPQPAGDDACSVQILVPGLKGDAGEKGDKGAPGRPGRVG PTGEKGDMGDKGQKGSVGRHGKIGPIGSKGEKGDSGDIGPPGPNGEPGLPCECSQLRKAIGE MDNQVSQLTSELKFIKNAVAGVRETESKIYLLVKEEKRYADAQLSCQGRGGTLSMPKDEAAN GLMAAYLAQAGLARVFIGINDLEKEGAFVYSDHSPMRTFNKWRSGEPNNAYDEEDCVEMVAS GGWNDVACHTTMYFMCEFDKENM

## FIGURE 253

AGTGACTGCAGCCTTCCTAGATCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGCAG CACCAGTGTGTGAGGGGAGCAGCCAGCCTTGCCAGATCCTTGATCCTGCCAGACCACC  ${\tt CAGCCCCGGCACAGAGCTGCTCCACAGGCACC} \underline{\textbf{ATG}} \\ \texttt{AGGATCATGCTGCTATTCACAGCCAT}$ TTCCTGGCGGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTC AAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAA GGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGA GGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTTCTTACCTTCAGTGAGGGTTCCTCGGCCC CTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACC  $\verb|TTTA| \textbf{TAA} \textbf{GACTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGT| \\$ ATCCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTC TTTCCCTGTCCCAATCCCCAGGTGCGCACGCTCCTGTTACCCTTTTCTCTTCCCTGTTCTTGT AACATTCTTGTGCTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAAACTGCA TAGTGAATATCCCCAACCCCAATGGGCATTGACTGTAGAATACCCTAGAGTTCCTGTAGTGT CCTACATTAAAAATATAATGTCTCTCTCTATTCCTCAACAATAAAGGATTTTTGCATATGAA 

## FIGURE 254

MRIMLLFTAILAFSLAQSFGAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGLLK ALSQASTDPKESTSPEKRDMHDFFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLGSTGK SSLGTEEQRPL

#### Important features:

#### Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 36-45

N-myristoylation site.

amino acids 33-39, 59-65

Amidation site.

amino acids 90-94

Leucine zipper pattern.

amino acids 43-65

Tachykinin family signature.

amino acids 86-92

### FIGURE 255

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGCTGTGCCCGCTGTGCCTGTGCC CGCGCTGTCGCCGCTGCTACCGCGTCTGCTGGACGCGGGAGACGCCAGCGAGCTGGTGATTG GAGCCCTGCGGAGAGCTCAAGCGCCCAGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTC  ${\tt CCATAGTTGCTGCAGGAGTGGAGCC}$ TGCTGTTCCTGGTCTGCGGATCCCAAGGCTACCTCCTGCCCAACGTCACTCTCTTAGAGGAG CTGCTCAGCAAATACCAGCACAACGAGTCTCACTCCCGGGTCCGCAGAGCCATCCCCAGGGA GGACAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGCCCAGGTGCAGCCTCAGGCCT CCAACATGGAGTACATGGTGAGCGCCGGGCTCCGGCCGCAGAGGCTGGCACCGGGGGTGGGGC TTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAATCGTCATGCCCTG AAACCT**TAG**ACTCCCGGGGTTAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAACTACAG GCATGCACCATGGTGCCCAGCTAGATTTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGT TGCCCAGGCTGGTCTTGAACTCCTAGGCTCAAGCAATCCTCCTGCCTCAGCCTCTCAAAGTG CTAGGATTATAGGCATGAGTCACCCTGTCTGGCTCTGGCTCTGTTCTTAACATTCTGCCAAA ACAACACGTGGGTTCCCTGTGCAGAGCCTGCCTCGTTGCCTTCATGTCACTCTTGGTAGC TCCACTGGGAACACAGCTCTCAGCCTTTCCCACCTGGAGGCAGAGTGGGGAGGGGCCCAGGG CTGGGCTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTC GTGAGATAAGTCGAGGCTGTGAAGGGCCCGGCACAGACTGACCTGCCTCCCCAACCCCTAGG CTTTGCTAACCGGGAAAGGAGCTAACGGTGACAGAAGACAGCCAAGGTCAACCCTCCCGGGT GATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGACCCCAAGCTCCAGTG TGGAAACTTCCTTGCTGGCTGGTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAGG GTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTCAGCTATGAATGGCTT TTTAAACAAACCCACGTCCCAGCCTGGGTAACATGGTAAAGCCCCGTCTCTACAAAAAAATC CAAGTTAGCCGGGCATGGTGGTGCGCACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAG GTGGAGGTGGGGGGGGGGGGGGAGGGAGGGATCGCTTGAGCCTGGGAAGTCGAGGCTGC AGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCAAAAA

# FIGURE 256

MSCVLGGVIPLGLLFLVCGSQGYLLPNVTLLEELLSKYQHNESHSRVRRAIPREDKEEILML HNKLRGQVQPQASNMEYMVSAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGWLRVSSGR GGSRLCSVLFVCFETGSHSATDAGVQWHNRHALKP

### Important features:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 27-31, 41-45

N-myristoylation site.

amino acids 126-132, 140-146

Amidation site.

amino acids 85-89

## FIGURE 257

# FIGURE 258

 ${\tt MGSGLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSGTS} \\ {\tt VTLHHARSQHHVVCNT}$ 

### FIGURE 259

# FIGURE 260

MIGYYLILFLMWGSSTVFCVLLIFTIAEASFSVENECLVDLCLLRICYKLSGVPNQCRVPLP SDCSK

Important features:

Signal peptide:

amino acids 1-29

### FIGURE 261

GAGGATTTGCCACAGCAGCGGATAGAGCAGGAGGAGCACCACCGGAGCCCTTGAGACATCCTT  ${\tt GAGAAGAGCCACAGCATAAGAGACTGCCCTGCTTGGTGTTTTGCAGG} {\color{red} {\bf ATG}} {\tt ATGGTGGCCCTT}$ CGAGGAGCTTCTGCATTGCTGGTTCTGTTCCTTGCAGCTTTTCTGCCCCCGCCGCAGTGTAC CCAGGACCCAGCCATGGTGCATTACATCTACCAGCGCTTTCGAGTCTTGGAGCAAGGGCTGG AAAAATGTACCCAAGCAACGAGGGCATACATTCAAGAATTCCAAGAGTTCTCAAAAAATATA TCTGTCATGCTGGGAAGATGTCAGACCTACACAAGTGAGTACAAGAGTGCAGTGGGTAACTT ACGAGTGCATCGTATCAGAGGACAAGACACTGGCAGAAATGTTGCTCCAAGAAGCTGAAGAA GAGAAAAAGATCCGGACTCTGCTGAATGCAAGCTGTGACAACATGCTGATGGGCATAAAGTC TTTGAAAATAGTGAAGAAGATGATGGACACACATGGCTCTTGGATGAAAGATGCTGTCTATA ACTCTCCAAAGGTGTACTTATTAATTGGATCCAGAAACAACACTGTTTGGGAATTTGCAAAC ATACGGGCATTCATGGAGGATAACACCAAGCCAGCTCCCCGGAAGCAAATCCTAACACTTTC CTGGCAGGGAACAGGCCAAGTGATCTACAAAGGTTTTCTATTTTTTCATAACCAAGCAACTT CTAATGAGATAATCAAATATAACCTGCAGAAGAGGGACTGTGGAAGATCGAATGCTGCTCCCA GGAGGGGTAGGCCGAGCATTGGTTTACCAGCACTCCCCCTCAACTTACATTGACCTGGCTGT GGATGAGCATGGGCCTCTGGGCCATCCACTCTGGGCCAGGCACCCATAGCCATTTGGTTCTCA CAAAGATTGAGCCGGGCACACTGGGAGTGGAGCATTCATGGGATACCCCATGCAGAAGCCAG GATGCTGAAGCCTCATTCCTCTTGTGTGGGGGTTCTCTATGTGGTCTACAGTACTGGGGGCCA GGGCCCTCATCGCATCACCTGCATCTATGATCCACTGGGCACTATCAGTGAGGAGGACTTGC CCAACTTGTTCTTCCCCAAGAGACCAAGAAGTCACTCCATGATCCATTACAACCCCAGAGAT AAGCAGCTCTATGCCTGGAATGAAGGAAACCAGATCATTTACAAACTCCAGACAAAGAGAAAA AGTGTGTAGAAGTGGAAATACGTATGCCTCCTTTCCCAAATGTCACTGCCTTAGGTATCTTC CAAGAGCTTAGATGAGAGCATATCATCAGGAAAGTTTCAACAATGTCCATTACTCCCCCAAA CCTCCTGGCTCTCAAGGATGACCACATTCTGATACAGCCTACTTCAAGCCTTTTGTTTTACT CCCTAATATTCACCACTGGCTTTTCTCTCCCCTGGCCTTTGCTGAAGCTCTTCCCTCTTTTT CAAATGTCTATTGATATTCTCCCATTTTCACTGCCCAACTAAAATACTATTAATATTTCTTT CTTTTCTTTTCTTTTTTTGAGACAAGGTCTCACTATGTTGCCCAGGCTGGTCTCAAACTCC AGAGCTCAAGAGATCCTCCTGCCTCAGCCTCCTAAGTACCTGGGATTACAGGCATGTGCCAC CACACCTGGCTTAAAATACTATTTCTTATTGAGGTTTAACCTCTATTTCCCCTAGCCCTGTC CTTCCACTAAGCTTGGTAGATGTAATAATAAAGTGAAAATATTAACATTTGAATATCGCTTT CCAGGTGTGGAGTGTTTGCACATCATTGAATTCTCGTTTCACCTTTGTGAAACATGCACAAG TCTTTACAGCTGTCATTCTAGAGTTTAGGTGAGTAACACAATTACAAAGTGAAAGATACAGC TAGAAAATACTACAAATCCCATAGTTTTTCCATTGCCCAAGGAAGCATCAAATACGTATGTT TGTTCACCTACTCTTATAGTCAATGCGTTCATCGTTTCAGCCTAAAAATAATAGTCTGTCCC TTTAGCCAGTTTTCATGTCTGCACAAGACCTTTCAATAGGCCTTTCAAATGATAATTCCTCC AGAAAACCAGTCTAAGGGTGAGGACCCCAACTCTAGCCTCCTCTTGTCTTGCTGTCCTCTGT 

## FIGURE 262

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFRVLEQGLEKCTQATRAYIQEFQE
FSKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLL
QEAEEEKKIRTLLNASCDNMLMGIKSLKIVKKMMDTHGSWMKDAVYNSPKVYLLIGSRNNTV
WEFANIRAFMEDNTKPAPRKQILTLSWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVED
RMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKIEPGTLGVEHSWDT
PCRSQDAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIH
YNPRDKQLYAWNEGNQIIYKLQTKRKLPLK

### FIGURE 263

GGGCGCCCGCGTACTCACTAGCTGAGGTGGCAGTGGTTCCACCAAC<u>ATG</u>GAGCTCTCGCAGA TGTCGGAGCTCATGGGGCTGTCGGTGTTGCTTGGGCTGCTGGCCCTGATGGCGACGGCGGCG AAATGGATTTCCACCTGACAAATCTTCGGGATCCAAGAAGCAGAAACAATATCAGCGGATTC GGAAGGAGAAGCCTCAACAACACAACTTCACCCACCGCCTCCTGGCTGCAGCTCTGAAGAGC CACAGCGGGAACATATCTTGCATGGACTTTAGCAGCAATGGCAAATACCTGGCTACCTGTGC AGATGATCGCACCATCCGCATCTGGAGCACCAAGGACTTCCTGCAGCGAGAGCACCGCAGCA TGAGAGCCAACGTGGAGCTGGACCACCCTGGTGCGCTTCAGCCCTGACTGCAGAGCC TTCATCGTCTGGCCGGCCAACGGGGACACCCTCCGTGTCTTCAAGATGACCAAGCGGGAGGA TGGGGGCTACACCTTCACAGCCACCCCAGAGGACTTCCCTAAAAAGCACAAGGCGCCTGTCA TCGACATTGGCATTGCTAACACAGGGAAGTTTATCATGACTGCCTCCAGTGACACCACTGTC ACACGCTGCTGTATCTCCCTGTGGCAGATTTGTAGCCTCGTGTGGCTTCACCCCAGATGTGA AGGTTTGGGAAGTCTGCTTTGGAAAGAAGGGGGGAGTTCCAGGAGGTGGTGCGAGCCTTCGAA CTAAAGGGCCACTCCGCGGCTGTGCACTCGTTTGCTTTCTCCAACGACTCACGGAGGATGGC TTCTGTCTCCAAGGATGGTACATGGAAACTGTGGGACACAGATGTGGAATACAAGAAGAAGC AGGACCCCTACTTGCTGAAGACAGGCCGCTTTGAAGAGGCGGCGGGTGCCGCCGTGCCGC CTGGCCCTCTCCCCCAACGCCCAGGTCTTGGCCTTGGCCAGTGGCAGTAGTATTCATCTCTA CAATACCCGGCGGGGCGAGAAGGAGGAGTGCTTTGAGCGGGTCCATGGCGAGTGTATCGCCA ACTTGTCCTTTGACATCACTGGCCGCTTTCTGGCCTCCTGTGGGGACCGGGCGGTGCGGCTG TTTCACAACACTCCTGGCCACCGAGCCATGGTGGAGGAGATGCAGGGCCACCTGAAGCGGGC CTCCAACGAGAGCACCCGCCAGAGGCTGCAGCAGCTGACCCAGGCCCAAGAGACCCTGA GGATCTGGCCTCCTCATGGCACTGCTGCCATCTTTCCTCCCAGGTGGAAGCCTTTCAGAAGG AGTCTCCTGGTTTTCTTACTGGTGGCCCTGCTTCTTCCCATTGAAACTACTCTTGTCTACTT AGGTCTCTCTTCTTGCTGGCTGTGACTCCTCCCTGACTAGTGGCCAAGGTGCTTTTCTTC CTCCCAGGCCCAGTGGGTGGAATCTGTCCCCACCTGGCACTGAGGAGAATGGTAGAGAGGAG AGGAGAGAGAGAGAATGTGATTTTTGGCCTTGTGGCAGCACATCCTCACACCCAAAGAAG TTTGTAAATGTTCCAGAACAACCTAGAGAACACCTGAGTACTAAGCAGCAGTTTTGCAAGGA TGGGAGACTGGGATAGCTTCCCATCACAGAACTGTGTTCCCATCAAAAAGACACTAAGGGATT TCCTTCTGGGCCTCAGTTCTATTTGTAAGATGGAGAATAATCCTCTCTGTGAACTCCTTGCA AAGATGATATGAGGCTAAGAGAATATCAAGTCCCCAGGTCTGGAAGAAAAGTAGAAAAGAGT AGTACTATTGTCCAATGTCATGAAAGTGGTAAAAGTGGGAACCAGTGTGCTTTGAAACCAAA TTAGAAACACATTCCTTGGGAAGGCAAAGTTTTCTGGGACTTGATCATACATTTTATATGGT TGGGACTTCTCTCTGGGAGATGATATCTTGTTTAAGGAGACCTCTTTTCAGTTCATCAAG 

### FIGURE 264

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQK
QYQRIRKEKPQQHNFTHRLLAAALKSHSGNISCMDFSSNGKYLATCADDRTIRIWSTKDFLQ
REHRSMRANVELDHATLVRFSPDCRAFIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPKK
HKAPVIDIGIANTGKFIMTASSDTTVLIWSLKGQVLSTINTNQMNNTHAAVSPCGRFVASCG
FTPDVKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTDV
EYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH
GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHLKRASNESTRQRLQQQLTQ
AQETLKSLGALKK

#### Important features:

### Signal peptide:

amino acids 1-25

#### N-glycosylation site.

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

### Beta-transducin family Trp-Asp repeat protein.

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

### FIGURE 265

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGAGGAAGCGAGAGGCATCTAAGCAGG CAGTGTTTTGCCTTCACCCCAAGTGACC**ATG**AGAGGTGCCACGCGAGTCTCAATCATGCTCC TCCTAGTAACTGTGTCTGACTGTGTGATCACAGGGGCCTGTGAGCGGGATGTCCAGTGT GGGGCAGCCTGCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCT GGGGCGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCA AGCACCACCTGTCCTTGCTTGCCCAACCTGCTGTGCTCCAGGTTCCCGGACGGCAGGTAC CGCTGCTCCATGGACTTGAAGAACATCAATTTT TAG GCGCTTGCCTGGTCTCAGGATACCCACCATCCTTTTCCTGAGCACAGCCTGGATTTTTATTTCTGCCATGAAACCCAGCTCCCATGAC GCAGACATACCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGG AATGGCAGAAAGGACATTCCCCCTCCCCTCCCCAGGTGACCTGCTCTCTTTCCTGGGCCCTG CCCCTCTCCCCACATGTATCCCTCGGTCTGAATTAGACATTCCTGGGCACAGGCTCTTGGGT GCATTGCTCAGAGTCCCAGGTCCTGGCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGG ACCAATTTGTGGGTAGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGAC TCAAGATTGGCTCTTCCCAGAGGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCCA GGGAGGCCAATCAGCCCCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGA CCTGTGACCTTCTGCCAGAATTGTCATGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGT TAACCACTGAAGCCCCCAATTCCCACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAA TCTAATCTGATATTGACATATTAGAAGGCAATTAGGGTGTTTCCTTAAACAACTCCTTTCCA AGGATCAGCCCTGAGAGCAGGTTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGG TGGGAGCAAGGGACAGGGGCAGGGGCTGAAAGGGGCACTGATTCAGACCAGGGAGG CAACTACACCAACATGCTGGCTTTAGAATAAAAGCACCAACTGAAAAAA

## FIGURE 266

 ${\tt MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAISLWLRGLRMCTPLGREGEECHP} \\ {\tt GSHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF}$ 

### Signal peptide:

amino acids 1-19

### Tyrosine kinase phosphorylation site:

amino acids 88-95

### N-myristoylation sites:

amino acids 33-39, 35-41, 46-52

### FIGURE 267

 ${\tt AGCGCCCGGGCGTCGGGGGGGTAAAAGGCCGGCAGAAGGGAGGCACTTGAGAA} \underline{{\tt ATG}} {\tt TCTTTC}$ TGCTGCCTTGGCATTGCTGCCAACACAGACGTGTTTCTGTCCAAGCCCCAGAAAGCGG CCCTGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTTCAAA GCAAAGGAGCTATGGGAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGCCAGGCTGTTT CCTCTGTCGAGAGGAAGCTGCGGATCTGTCCTCCCTGAAAAGCATGTTGGACCAGCTGGGCG TCCCCCTCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGGATTTCCAGCCTTAT TTCAAAGGAGAAATCTTCCTGGATGAAAAGAAAAAGTTCTATGGTCCACAAAGGCGGAAGAT GATGTTTATGGGATTTATCCGTCTGGGAGTGTGGTACAACTTCTTCCGAGCCTGGAACGGAG GCTTCTCTGGAAACCTGGAAGGAGAAGGCTTCATCCTTGGGGGAGTTTTCGTGGTGGGATCA GGAAAGCAGGGCATTCTTCTTGAGCACCGAGAAAAAGAATTTGGAGACAAAGTAAACCTACT TTCTGTTCTGGAAGCTGCTAAGATGATCAAACCACAGACTTTGGCCTCAGAGAAAAAA ${f TGA}$ T TGTGTGAAACTGCCCAGCTCAGGGATAACCAGGGACATTCACCTGTGTTCATGGGATGTATT GTTTCCACTCGTGTCCCTAAGGAGTGAGAAACCCATTTATACTCTACTCTCAGTATGGATTA TTAATGTATTTTAATATTCTGTTTAGGCCCACTAAGGCAAAATAGCCCCAAAACAAGACTGA CAAAAATCTGAAAAACTAATGAGGATTATTAAGCTAAAACCTGGGAAATAGGAGGCTTAAAA TTGACTGCCAGGCTGGGTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGG TGAGCAAGTCACTTGAGGTCGGGAGTTCGAGACCAGCCTGAGCAACATGGCGAAACCCCGTC GGAGGCTGAGGCAGGAGATCACTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCA 

# FIGURE 268

MSFLQDPSFFTMGMWSIGAGALGAAALALLLANTDVFLSKPQKAALEYLEDIDLKTLEKEPR TFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDF QPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGVWYNFFRAWNGGFSGNLEGEGFILGGVFV VGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

### FIGURE 269

# FIGURE 270

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSSTSSSSDGNLRPEAITAIIV VFSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

### FIGURE 271

# FIGURE 272

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGK GIVKGRNLDSRGLILGAEAWGRGVKKNT

## FIGURE 273

 ${\tt GCCAGGAATAACTAGAGAGGAACA} {\color{red} {\bf ATG}} {\tt GGGTTATTCAGAGGTTTTGTTTTCCTCTTAGTTCT}$ GTGCCTGCTGCACCAGTCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATA TTGTCATTGTTATAGATCCTAGTGTGCCAGAAGATGAAAAAATAATTGAACAAATAGAGGAT ATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAAGATTTTTTTCAAAAA TGTATCTATATTAATTCCTGAGAATTGGAAGGAAAATCCTCAGTACAAAAGGCCAAAACATG AAAACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCA TACACCAAGCAGTTCACAGAATGTGGAGAGAAAGGCGAATACATTCACTTCACCCCTGACCT TCTACTTGGAAAAAAAAATGAATATGGACCACCAGGCAAACTGTTTGTCCATGAGTGGG CTCACCTCCGGTGGGGAGTGTTTGATGAGTACAATGAAGATCAGCCTTTCTACCGTGCTAAG TCAAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAA GTGTCAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACTGTATG GAAAAGATTGTCAATTCTTTCCTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATG CAAAGTATTGATTCTGTTGTTGAATTTTGTAACGAAAAAACCCATAATCAAGAAGCTCCAAG CCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAATTCTGAGGATT AGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGAACTTCCAT CTGCTCTGGAATTAAATATGCATTTCAGGTGATTGGAGAGCTACATTCCCAACTCGATGGAT CCGAAGTACTGCTGACTGATGGGGAGGATAACACTGCAAGTTCTTGTATTGATGAAGTG AAACAAAGTGGGGCCATTGTTCATTTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAAT AGAGATGAGCAAGATAACAGGAGGAAGTCATTTTTATGTTTCAGATGAAGCTCAGAACAATG GCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCCAGAAGTCCCTT CAGCTCGAAAGTAAGGGATTAACACTGAATAGTAATGCCTGGATGAACGACACTGTCATAAT TGATAGTACAGTGGGAAAGGACACGTTCTTTCTCATCACATGGAACAGTCTGCCTCCCAGTA TTTCTCTCTGGGATCCCAGTGGAACAATAATGGAAAATTTCACAGTGGATGCAACTTCCAAA ATGGCCTATCTCAGTATTCCAGGAACTGCAAAGGTGGGCACTTGGGCATACAATCTTCAAGC CAAAGCGAACCCAGAAACATTAACTATTACAGTAACTTCTCGAGCAGCAAATTCTTCTGTGC CTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCCAGCCCAATGATT GTTTACGCAGAAATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTCAT TGAATCACAGAATGGACATACAGAAGTTTTGGAACTTTTGGATAATGGTGCAGGCGCTGATT CTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATATACAGAAAATGGCAGATAT AGCTTAAAAGTTCGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAAATTACGGCCTCCACT GAATAGAGCCGCGTACATACCAGGCTGGGTAGTGAACGGGGAAATTGAAGCAAACCCGCCAA GACCTGAAATTGATGAGGATACTCAGACCACCTTGGAGGATTTCAGCCGAACAGCATCCGGA GGTGCATTTGTGGTATCACAAGTCCCAAGCCTTCCCTTGCCTGACCAATACCCACCAAGTCA AATCACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTCTTACATGGACAGCACCAG GAGATAATTTTGATGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT GATCTAAGAGACAGTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTCACCAAAGGA GGCCAACTCCAAGGAAAGCTTTGCATTTAAACCAGAAAATATCTCAGAAGAAAATGCAACCC ACATATTTATTGCCATTAAAAGTATAGATAAAAGCAATTTGACATCAAAAGTATCCAACATT GCACAAGTAACTTTGTTTATCCCTCAAGCAAATCCTGATGACATTGATCCTACACCTACTCC
TACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTTCTACGCTGGTAT
TGTCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTTAAGTACCACCATT**TGA**ACCTTA AAAGGATATTTCTGAATCTTAAAATTCATCCCATGTGTGATCATAAACTCATAAAAATAATT TTAAGATGTCGGAAAAGGATACTTTGATTAAATAAAAACACTCATGGATATGTAAAAACTGT CAAGATTAAAATTTAATAGTTTCATTTATTTGTTATTTTGTAAGAAATAGTGATGAAC AAAGATCCTTTTTCATACTGATACCTGGTTGTATATTATTTGATGCAACAGTTTTCTGAAAT 

### FIGURE 274

MGLFRGFVFLLVLCLLHQSNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTTASTY
LFEATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTEC
GEKGEYIHFTPDLLLGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATR
CSAGISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVVE
FCNEKTHNQEAPSLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLLKISQRIVCLV
LDKSGSMGGKDRLNRMNQAAKHFLLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLM
AGLPTYPLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLTDGEDNTASSCIDEVKQSGAIVH
FIALGRAADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLT
LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPG
TAKVGTWAYNLQAKANPETLTITVTSRAANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQG
YVPVLGANVTAFIESQNGHTEVLELLDNGAGADSFKNDGVYSRYFTAYTENGRYSLKVRAHG
GANTARLKLRPPLNRAAYIPGWVVNGEIEANPPRPEIDEDTQTTLEDFSRTASGGAFVVSQV
PSLPLPDQYPPSQITDLDATVHEDKIILTWTAPGDNFDVGKVQRYIIRISASILDLRDSFDD
ALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAIKSIDKSNLTSKVSNIAQVTLFIP
OANPDDIDPTPTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

#### Signal peptide:

amino acids 1-21

#### Putative transmembrane domains:

amino acids 284-300, 617-633

#### Leucine zipper pattern.

amino acids 469-491, 476-498

#### N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592, 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

## FIGURE 275

CCCACGGGGAAGGAGGGTACCAGGCCGTCCTTCAGGAGTGGGAGCAGCTGCCCAGGGCCAACAGC GCAGCCTGAAGCGCCAGATCGCACAGCTCAAGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGAACTACGTGA 

### FIGURE 276

MMMVRRGLLAWISRVVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQ
EWEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFL
HSQVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIES
ALETLNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPI
MKVKNEKLNMANTLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVK
GILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCR
LNTQPGKKVFYPVLFSQYNPGIIYGHHDAVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFI
NIGGFDLDIKGWGGEDVHLYRKYLHSNLIVVRTPVRGLFHLWHEKRCMDELTPEQYKMCMQS
KAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT

### FIGURE 277

GAAAGA**ATG**TTGTGGCTGCTCTTTTTTCTGGTGACTGCCATTCATGCTGAACTCTGTCAACC AGGTGCAGAAAATGCTTTTAAAGTGAGACTTAGTATCAGAACAGCTCTGGGAGATAAAGCAT ATGCCTGGGATACCAATGAAGAATACCTCTTCAAAGCGATGGTAGCTTTCTCCATGAGAAAA GTTCCCAACAGAAAGCAACAGAAATTTCCCATGTCCTACTTTGCAATGTAACCCAGAGGGT ATCATTCTGGTTTGTGGTTACAGACCCTTCAAAAAATCACACCCTTCCTGCTGTTGAGGTGC CTGGAATTTTTAAAAATCCCTTCCACACTTGCACCACCCATGGACCCATCTGTGCCCATCTG GATTATTATATTTGGTGTGATATTTTGCATCATCATAGTTGCAATTGCACTACTGATTTTAT CAGGGATCTGGCAACGTAGAAGAAGAACAAAGAACCATCTGAAGTGGATGACGCTGAAGAT AAGTGTGAAAACATGATCACAATTGAAAATGGCATCCCCTCTGATCCCCTGGACATGAAGGG GGGCATATTAATGATGCCTTCA**TGA**CAGAGGATGAGAGGCTCACCCCTCTCTGAAGGGCTGT TGTTCTGCTTCCTCAAGAAATTAAACATTTGTTTCTGTGTGACTGCTGAGCATCCTGAAATA CCAAGAGCAGATCATATATTTTGTTTCACCATTCTTCTTTTTGTAATAAATTTTGAATGTGCT TGAAAGTGAAAAGCAATCAATTATACCCACCAACACCCCTGAAATCATAAGCTATTCACGAC TCAAAATATTCTAAAATATTTTTCTGACAGTATAGTGTATAAATGTGGTCATGTGGTATTTG TAGTTATTGATTTAAGCATTTTTAGAAATAAGATCAGGCATATGTATATATTTTCACACTTC AAAGACCTAAGGAAAAATAAATTTTCCAGTGGAGAATACATATAATATGGTGTAGAAATCAT TGAAAATGGATCCTTTTTGACGATCACTTATATCACTCTGTATATGACTAAGTAAACAAAAG TGAGAAGTAATTATTGTAAATGGATGGATAAAAATGGAATTACTCATATACAGGGTGGAATT TTATCCTGTTATCACACCAACAGTTGATTATATTTTTCTGAATATCAGCCCCTAATAGGAC AATTCTATTTGTTGACCATTTCTACAATTTGTAAAAGTCCAATCTGTGCTAACTTAATAAAG 

# FIGURE 278

MLWLLFFLVTAIHAELCQPGAENAFKVRLSIRTALGDKAYAWDTNEEYLFKAMVÁFSMRKVP NREATEISHVLLCNVTQRVSFWFVVTDPSKNHTLPAVEVQSAIRMNKNRINNAFFLNDQTLE FLKIPSTLAPPMDPSVPIWIIIFGVIFCIIIVAIALLILSGIWQRRRKNKEPSEVDDAEDKC ENMITIENGIPSDPLDMKGGILMMPS

### FIGURE 279

AACTCAAACTCCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGG GTGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC**ATG**T ATGGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATACAGCTCACAGCTCTTTGG CCTATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGC TCGGTTAAAATGCACTTTCTCCAGCTTTGCCCCTGTGGGTGATGCTCTAACAGTGACCTGGA ATTTTCGTCCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTC CAACCCATGAGTGGGCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGA TGCCTCCATCCTTCTCTGGAAACTGCAGTTCGACGACAATGGGACATACACCTGCCAGGTGA AGAACCCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTA CGCTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCAT AATAGTAATTGTAGTGGTCCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTC ATAAAGTGGTGGAGATAAAATCAAAAGAAGAGGGAAAGGCTCAACCAAGAGAAAAAGGTCTCT  $\tt GTTTATTTAGAAGACACAGAC{\bf TAA}{\tt CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAA$ GAACCCTAGTATTTCTTGAAGTTAATGGAAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGT TTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCAC AGTGCTCCTCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTCAGA GTGTAAATTTTTTCAAGTGCTCATTAGGTTTTATAAACAAGAAGCTACATTTTTGCCCCTTAA GACACTACTTACAGTGTTATGACTTGTATACACATATATTGGTATCAAAGGGGGATAAAAGCC AATTTGTCTGTTACATTTCCTTTCACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTA ATGTGTTTACTCTCTTCCCTTCCCACATTCTCAATTAAAAGGTGAGCTAAGCCTCCTCGGTG TTTGTCG

# FIGURE 280

MYGKSSTRAVLLLLGIQLTALWPIAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVT WNFRPLDGGPEQFVFYYHIDPFQPMSGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQ VKNPPDVDGVIGEIRLSVVHTVRFSEIHFLALAIGSACALMIIIVIVVVLFQHYRKKRWAER AHKVVEIKSKEEERLNQEKKVSVYLEDTD

### FIGURE 281

# FIGURE 282

 ${\tt MKFLAVLVLLGVSIFLVSAQNPTTAAPADTYPATGPADDEAPDAETTAAATTATTAAPTTAT}$   ${\tt TAASTTARKDIPVLPKWVGDLPNGRVCP}$ 

### FIGURE 283

GGACTCTGAAGGTCCCAAGCAGCTGCTGAGGCCCCCAAGGAAGTGGTTCCAACCTTGGACCC
CTAGGGGTCTGGATTTGCTGGTTAACAAGATAACCTGAGGGCAGGACCCCATAGGGGAATGC
TACCTCCTGCCCTTCCACCTGCCCTGGTGTTCACGGTGGCCTGGTCCCTCCTTGCCGAGAGA
GTGTCCTGGGTCAGGGACGCAGAGGACGCTCACAGACTCCAGCCCTTTGTTACCGAGAGAC
ACTTGGCAAGGTCCAGCGATGGTCCGGAGTCCACACACAGACTGGCGGCAGGGCAGGAGGGC
GACAGTTCTGTTGTGCTTGGTTGGACAGTAAGAGGGTCTTGGCCAGTCCAGGGTGGGGGGG
GCAAACTCCATAAAGAACCAGAGGGTCTGGGCCCCGGCCACAGAGTCATCTGCCCAGCTCCT
CTGCTGCTGGCCAGTGGGAGTGGCACGAGGTGGGGGCAGAGTCATCTGCCCAGCTCCT
CTGCTGCTGCCGGCCATGGTCCCTGTCTAGGGCAGCAATTCTCAACCTTCTTGCTCTCAGGA
ATTTGCCTGCGGGCCATGGTCCCTGTCTAGGGCAGCAATTCTCAACCTTCTTGCTCTCAGGA
CCCCAAAGAGCTTTCATTGTATCTATTGATTTTTACCACATTAGCAATTAAAACCTGAGAAAT
GGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGAT
CACCTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTACTAAAAA
TACAAAAAATTAGCCAGGCACAGTGGTGTGCACTGGTAGTCCCAGTTACTCGGGAGGCTGAG
GCAGGAAAATCGCTTGAACCCAGGAGGCGGACGTTGCGGTGAGCCGAGATCGCCCCCTGAT
TCCAGCCTGGGCGACAAAGAGTGAGACTCCATCTCACACA

# FIGURE 284

MLPPALPPALVFTVAWSLLAERVSWVRDAEDAHRLQPFVTERTLGKVQRWSGVHTQTGGRAG GGQFCCAWLDSKRVLASPGWGAANSIKNQRVWAPATESSAQLLCCWPVGVARGGALCQ

### FIGURE 285

## FIGURE 286

MPVPALCLLWALAMVTRPASAAPMGGPELAQHEELTLLFHGTLQLGQALNGVYRTTEGRLTK ARNSLGLYGRTIELLGQEVSRGRDAAQELRASLLETQMEEDILQLQAEATAEVLGEVAQAQK VLRDSVQRLEVQLRSAWLGPAYREFEVLKAHADKQSHILWALTGHVQRQRREMVAQQHRLRQ IQERLHTAALPA

## FIGURE 287

GGCAAC**ATG**GCTCAGCAGGCTTGCCCCAGAGCCATGGCAAAGAATGGACTTGTAATTTGCAT CCTGGTGATCACCTTACTCCTGGACCAGACCACCAGCCACACATCCAGATTAAAAGCCAGGA AGCACAGCAAACGTCGAGTGAGAGACAAGGATGGAGATCTGAAGACTCAAATTGAAAAGCTC TAAAGTTCACAAGAAATGCTACCTTGCTTCAGAAGGTTTGAAGCATTTCCATGAGGCCAATG AAGACTGCATTTCCAAAGGAGGAATCCTGGTTATCCCCAGGAACTCCGACGAAATCAACGCC CTCCAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGACTTTTGGCTGGGCATCAATGA ACCGTGCACAGCCTAACGGTGGCAAGCGAGAAAACTGTGTCCTGTTCTCCCAATCAGCTCAG GGCAAGTGGAGTGATGAGGCCTGTCGCAGCAGCAAGAGATACATATGCGAGTTCACCATCCC TAAA**TAG**GTCTTTCTCCAATGTGTCCTCCAAGCAAGATTCATCATAACTTATAGGTTCATGA TCTCTAAGATCAAGTAAAAATCATAATTTTTACTTATTAAAAAATTGCAACACAAGATCAAT GTCCATAGCAATATGATAGCATCAGCCAATTTTGCTAACACATTTCTTTGGGATTTTGCCCT TCCTGGGGTATAGGGGATCAGAAATATTGATCCATGTGCACGCAGATAAAATGGCTTCTGCT TTCCCTTGGAAGTTTAGCGTATGTTTGACTAACAAAAATTCCCTACATCAGAGACTCTAGGT GCTATATAATCCAAAAACTTTTCAGCCTGTTGCTCATTCTGTCCCATGCTGGCAATAATACC TTGTCAGCCCATTACCCTTATTTTGAATTGCTCCATCTCCTGGTGGGACTTGTATCTTGTCT TACCCTTTTTTTGGAAGTTTCCAGCCGCAATTTGAAATGAAATGACAAGGTGTATATTTGAT CAATTTTCATTCCCACCATTGCATTACAACCTCTAACTTAAATGGGTAACCCTAAGGCATAT AGCATCCTTACTCTCACCTTTTATGAGATTGAGAGTGGACTTACATTTCCTTTTTTACATTT TCGTATATTTATTTTTTTTAGCCATCATTATATGTTTAAGTCTATTATGGGCAACCAATCTT TGGAAGCTGAAAACTGAATTTAAAGAATGCTATCTTGGAAAATTGCATACGTCTGTGCAATT TTTTATTCTGCCTAGTGCTATTCTGCTTGTTTAACTAGATTGTACAAAATAACTTCATTGCT TAATATCAAATTACAAAGTTTAGACTTGGAGGGAAATGGGCTTTTTAGAAGCAAACAATTTT AAATATATTTTGTTCTTCAAATAAATAGTGTTTAAACATTGAATGTGTTTTGTGAACAATAT CCCACTTTGCAAACTTTAACTACACATGCTTGGAATTAAGTTTTAGCTGTTTTCATTGCTCA 

# FIGURE 288

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHSKRRVRDKDGDLKTQIEKLWT EVNALKEIQALQTVCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSQSAQGK WSDEACRSSKRYICEFTIPK

# FIGURE 289

# FIGURE 290

 ${\tt MKLAALLGLCVALSCSSAAAFLVGSAKPVAQPVAALESAAEAGAGTLANPLGTLNPLKLLLS} \\ {\tt SLGIPVNHLIEGSQKCVAELGPQAVGAVKALKALLGALTVFG}$ 

### FIGURE 291

TGAAGGACTTTTCCAGGACCCAAGGCCACACACTGGAAGTCTTGCAGCTGAAGGGAGGCACT CCTTGGCCTCCGCAGCCGATCAC<u>ATG</u>AAGGTGGTGCCAAGTCTCCTGCTCTCCGTCCTCCTG GCACAGGTGTGGCTGGTACCCGGCTTGGCCCCCAGTCCTCAGTCGCCAGAGACCCCCAGCCCC TCAGAACCAGACCAGCGGTAGTGCAGGCTCCCAGGGAGGAAGAGGAAGATGAGCAGGAGG CCAGCGAGGAGAAGGCCGGTGAGGAAGAGAAAGCCTGGCTGATGGCCAGCAGCAGCAGCTT GCCAAGGAGACTTCAAACTTCGGATTCAGCCTGCTGCGAAAGATCTCCATGAGGCACGATGG CAACATGGTCTTCTCTCCATTTGGCATGTCCTTGGCCATGACAGGCTTGATGCTGGGGGCCA CAGGGCCGACTGAAACCCAGATCAAGAGAGGGCTCCACTTGCAGGCCCTGAAGCCCACCAAG CCCGGGCTCCTGCCTTTTAAGGGACTCAGAGAGACCCTCTCCCGCAACCTGGAACT GGGCCTCTCACAGGGGAGTTTTGCCTTCATCCACAAGGATTTTGATGTCAAAGAGACTTTCT TCAATTTATCCAAGAGGTATTTTGATACAGAGTGCGTGCCTATGAATTTTCGCAATGCCTCA CAGGCCAAAAGGCTCATGAATCATTACATTAACAAAGAGACTCGGGGGAAAATTCCCAAACT GTTTGATGAGATTAATCCTGAAACCAAATTAATTCTTGTGGATTACATCTTGTTCAAAGGGA AATGGTTGACCCCATTTGACCCTGTCTTCACCGAAGTCGACACTTTCCACCTGGACAAGTAC AAGACCATTAAGGTGCCCATGATGTACGGTGCAGGCAAGTTTGCCTCCACCTTTGACAAGAA TTTTCGTTGTCATGTCCTCAAACTGCCCTACCAAGGAAATGCCACCATGCTGGTGGTCCTCA TGGAGAAAATGGGTGACCACCTCGCCCTTGAAGACTACCTGACCACAGACTTGGTGGAGACA GAAGTATGAGATGCATGAGCTGCTTAGGCAGATGGGAATCAGAAGAATCTTCTCACCCTTTG CTGACCTTAGTGAACTCTCAGCTACTGGAAGAAATCTCCAAGTATCCAGGGTTTTACGAAGA ACAGTGATTGAAGTTGATGAAAGGGGCACTGAGGCAGTGGCAGGAATCTTGTCAGAAATTAC AAACCTCIGGAATGCTTCTGTTTCTGGGCAGGGTGGTGAATCCGACTCTCCTA**TAA**TTCAGG TACCAGCAATGGATGGCAGGGGAGAGTGTTCCTTTTGTTCTTAACTAGTTTAGGGTGTTCTC AAATAAATACAGTAGTCCCCACTTATCTGAGGGGGATACATTCAAAGACCCCCAGCAGATGC AAAGTTTAATTTATAAATTAGGCACAGTAAGAGATTAACAATAATAACAACATTAAGTAAAA TGAGTTACTTGAACGCAAGCACTGCAATACCATAACAGTCAAACTGATTATAGAGAAGGCTA CTAAGTGACTCATGGGCGAGGAGCATAGACAGTGTGGAGACATTGGGCAAGGGGAGAATTCA CATCCTGGGTGGGACAGAGCAGGACGATGCAAGATTCCATCCCACTACTCAGAATGGCATGC TGCTTAAGACTTTTAGATTGTTTATTTCTGGAATTTTTCATTTAATGTTTTTTGGACCATGGT TGACCATGGTTAACTGAGACTGCAGAAAGCAAAACCATGGATAAGGGAGGACTACTACAAAA 

# FIGURE 292

MKVVPSLLLSVLLAQVWLVPGLAPSPQSPETPAPQNQTSRVVQAPREEEEDEQEASEEKAGE
EEKAWLMASRQQLAKETSNFGFSLLRKISMRHDGNMVFSPFGMSLAMTGLMLGATGPTETQI
KRGLHLQALKPTKPGLLPSLFKGLRETLSRNLELGLSQGSFAFIHKDFDVKETFFNLSKRYF
DTECVPMNFRNASQAKRLMNHYINKETRGKIPKLFDEINPETKLILVDYILFKGKWLTPFDP
VFTEVDTFHLDKYKTIKVPMMYGAGKFASTFDKNFRCHVLKLPYQGNATMLVVLMEKMGDHL
ALEDYLTTDLVETWLRNMKTRNMEVFFPKFKLDQKYEMHELLRQMGIRRIFSPFADLSELSA
TGRNLQVSRVLRRTVIEVDERGTEAVAGILSEITAYSMPPVIKVDRPFHFMIYEETSGMLLF
LGRVVNPTLL

# FIGURE 293

# FIGURE 294

MRRLLLVTSLVVVLLWEAGAVPAPKVPIKMQVKHWPSEQDPEKAWGARVVEPPEKDDQLVVL FPVQKPKLLTTEEKPRGQGRGPILPGTKAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQGEE RPRLWVMPNHQVLLGPEEDQDHIYHPQ

## FIGURE 295

TACCCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACA**ATG**AACCAACTCAGCTTCCTGC TGGACCTGTTCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCC TAGTGCATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTTATCTACCAGACCTTCT GTGACATGACCTCTGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATG CGTGGGAAGTGCACGGTGGCCGATCGCTGGTCCAGTCAGCAGGGCAAAGCAGACTACCC AGAGGGGGACGGCAACTGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCG ATGACTACAAGAACCCTGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTG CCCAATAAGTCCCCCATGCAGCACTGGAGAAACAGCTCCCTGCTGAGGTACCGCACGGACAC TGGCTTCCTCCAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCAGTGAAAT ATGGAGAAGGAAAGTGTTGGACTGACAACGGCCCGGTGATCCCTGTGGTCTATGATTTTGGC GACGCCCAGAAAACAGCATCTTATTACTCACCCTATGGCCAGCGGGAATTCACTGCGGGATT TGTTCAGTTCAGGGTATTTAATAACGAGAGAGCCAGCCAACGCCTTGTGTGCTGGAATGAGGG TCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAGGATACTTTCCAGAGGCCAGT CCCCAGCAGTGTGGAGATTTTTCTGGTTTTGATTGGAGTGGATATGGAACTCATGTTGGTTA  ${\sf CAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCGT}$ GGAGGGAACCCAGACCTCTCCCCAACCATGAGATCCCAAGGATGGAGAACAACTTACCCA GTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGAAAAAAA

# FIGURE 296

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTEN
GVIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFG
SAEAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGI
YQKYPVKYGEGKCWTDNGPVIPVVYDFGDAQKTASYYSPYGQREFTAGFVQFRVFNNERAAN
ALCAGMRVTGCNTEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLL
FYR

## FIGURE 297

GCGGAGCCGGCCGGCTGCGCAGAGGAGCCGCTCTCGCCGCCGCCACCTCGGCTGGGAGCC  ${\tt CACGAGGCTGCCGCATCCTGCCCTCGGAACA} {\color{red} {\bf ATG}} {\tt GGACTCGGCGCGCGAGGTGCTTGGGCCG}$ CGCTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGCCCCCATGAAAGCGCAGCC ATGGCGGCATCTGCAAACATAGAGAATTCTGGGCTTCCACACAACTCCAGTGCTAACTCAAC AGAGACTCTCCAACATGTGCCTTCTGACCATACAAATGAAACTTCCAACAGTACTGTGAAAC CACCAACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCACCACCATGAAACCTACAGCG GCATCTAATACAACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTC TACACCCAAAACAACAAGTGTTTCACAGAACACATCTCAGATATCAACATCCACAATGACCG TAACCCACAATAGTTCAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACTATGCAT TCTGAAGCAAAGAAAGGATCAAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAAC GCTGGGAGTTTTATCTATTCTTTACATTGGATGCAAAATGTATTACTCAAGAAGAGGCATTC GGTATCGAACCATAGATGAACATGATGCCATCATT**TAA**GGAAATCCATGGACCAAGGATGGA ATACAGATTGATGCTGCCCTATCAATTAATTTTGGTTTATTAATAGTTTAAAAACAATATTCT CTTTTTGAAAATAGTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTA AAGATTCTTCAAGGTAACAAGGGTTTGGGTTTTGAAATAAACATCTGGATCTTATAGACCGT GGGGTGGGGCATTGGTCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAA TGCCATCTGGGCATACAAATAAGAAGTTTGTCACAGCACTCAGGATTTTGGGTATCTTTTGT AGCTCACATAAAGAACTTCAGTGCTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACA CAGAAATTATACAATCAAACTAGATCTGAAGCATAATTTAAGAAAAACATCAACATTTTTTG TGCTTTAAACTGTAGTAGTTGGTCTAGAAACAAAATACTCC

# FIGURE 298

MGLGARGAWAALLLGTLQVLALLGAAHESAAMAASANIENSGLPHNSSANSTETLQHVPSDH TNETSNSTVKPPTSVASDSSNTTVTTMKPTAASNTTTPGMVSTNMTSTTLKSTPKTTSVSQN TSQISTSTMTVTHNSSVTSAASSVTITTTMHSEAKKGSKFDTGSFVGGIVLTLGVLSILYIG CKMYYSRRGIRYRTIDEHDAII

# FIGURE 299

CAGCCGGGTCCCAAGCCTGTGCCTGAGCCTGAGCCTGAGCCCGAGCCGGGAGCCGG CTTCTTCGGCTTGGCTCCTTGCTCAGCCTGTCGTGCCTGGCGCTTTCCGTGCTGCTGCTGGC ATAAAGAAAATTCTGGGCATATTTATAATAAGAACATATCTCAGAAAGATTGTGATTGCCTT CATGTTGTGGAGCCCATGCCTGTGCGGGGGCCTGATGTAGAAGCATACTGTCTACGCTGTGA ATGCAAATATGAAGAAGAAGCTCTGTCACAATCAAGGTTACCATTATAATTTATCTCTCCA TTTTGGGCCTTCTACTTCTGTACATGGTATATCTTACTCTGGTTGAGCCCATACTGAAGAGG CGCCTCTTTGGACATGCACAGTTGATACAGAGTGATGATGATATTGGGGATCACCAGCCTTT TGCAAATGCACACGATGTGCTAGCCCGCTCCCGCAGTCGAGCCAACGTGCTGAACAAGGTAG  ${\tt CATGTTGTCCTCAGC} \textbf{\underline{TAA}} {\tt TTGGGAATTGAATTCAAGGTGACTAGAAAGAAACAGGCAGACAA}$ CTGGAAAGAACTGACTGGGTTTTGCTGGGTTTCATTTTAATACCTTGTTGATTTCACCAACT ATAATAGAGACATTTTTAAAAGCACACAGCTCAAAGTCAGCCAATAAGTCTTTTCCTATTTG TGACTTTTACTAATAAAATAAATCTGCCTGTAAATTATCTTGAAGTCCTTTACCTGGAACA AGCACTCTCTTTTTCACCACATAGTTTTAACTTGACTTTCAAGATAATTTTCAGGGTTTTTTG TTGTTGTTTTTTTTTTTTTTTTTTTTTTTTGTTGGTGGGAGGGGAGGGATGCCTGGGAAGTGGTT AACAACTTTTTTCAAGTCACTTTACTAAACAAACTTTTGTAAATAGACCTTACCTTCTATTT TCGAGTTTCATTTATATTTTGCAGTGTAGCCAGCCTCATCAAAGAGCTGACTTACTCATTTG ACTTTTGCACTGACTGTATTATCTGGGTATCTGCTGTGTCTGCACTTCATGGTAAACGGGAT CTAAAATGCCTGGTGGCTTTTCACAAAAAGCAGATTTTCTTCATGTACTGTGATGTCTGATG CAATGCATCCTAGAACAAACTGGCCATTTGCTAGTTTACTCTAAAGACTAAACATAGTCTTG GTGTGTGTGGTCTTACTCATCTTCTAGTACCTTTAAGGACAAATCCTAAGGACTTGGACACT TGCAATAAAGAAATTTTATTTTAAACCCAAGCCTCCCTGGATTGATAATATACACATTTG TCAGCATTTCCGGTCGTGGTGAGAGGCAGCTGTTTGAGCTCCAATATGTGCAGCTTTGAACT AGGGCTGGGGTTGTGGGTGCCTCTTCTGAAAGGTCTAACCATTATTGGATAACTGGCTTTTT TCTTCCTATGTCCTCTTTGGAATGTAACAATAAAAATAATTTTTTGAAACATCAA

# FIGURE 300

MATLWGGLLRLGSLLSLSCLALSVLLLAQLSDAAKNFEDVRCKCICPPYKENSGHIYNKNIS QKDCDCLHVVEPMPVRGPDVEAYCLRCECKYEERSSVTIKVTIIIYLSILGLLLLYMVYLTL VEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDVLARSRSRANVLNKVEYAQQRWKLQVQEQ RKSVFDRHVVLS

## FIGURE 301

# FIGURE 302

MAYSTVQRVALASGLVLALSLLLPKAFLSRGKRQEPPPTPEGKLGRFPPMMHHHQAPSDGQT PGARFQRSHLAEAFAKAKGSGGGAGGGGSGRGLMGQIIPIYGFGIFLYILYILFKVSRIILI ILHQ

# FIGURE 303

 $\tt CGGCTCGAGTGCAGCTGTGGGGAGATTTCAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTT$  ${\tt GGATTTGAAAGTTGAGAGCAGC} \underline{\textbf{ATG}} \\ {\tt TTTTGCCCACTGAAACTCATCCTGCCAGTGTTAC}$ TGGATTATTCCTTGGGCCTGAATGACTTGAATGTTTCCCCGCCTGAGCTAACAGTCCATGTG GGTGATTCAGCTCTGATGGGATGTGTTTTCCAGAGCACAGAAGACAAATGTATATTCAAGAT AGACTGGACTCTGTCACCAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCA ATCTCAGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGACATCTTATGC AATGATGGCTCTCCTCCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGA AATCCGCCTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAG AGCACAGAAGTGAAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGA GGAGATTGTATTTCGTTACTACCACAAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGG GCCACTTCCAGAATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCCATCATG CTTCAAGGAGTGAGGGAGTCAGATGGAGGAAACTACACCTGCAGTATCCACCTAGGGAACCT GGTGTTCAAGAAAACCATTGTGCTGCATGTCAGCCCGGAAGAGCCCTCGAACACTGGTGACCC CGGCAGCCCTGAGGCCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTGTGGGAATTGTC TGTGCCACAATCCTGCTGCTCCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAA GAGTTCAGTGAATTCTACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAG AAAAACCCTGCCATTTTGAAAGATGTGAAGGGGAGAAACACATTTACTCCCCAATAATTGTA CGGGAGGTGATCGAGGAAGAAGAACCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCA CCCAGTTTGGCCTTCTCTGAGGTCAGATCGGAACAACTCACTTGAAAAAAAGTCAGGTGGGG GAATGCCAAAAACACAGCAAGCCTTT**TGA**GAAGAATGGAGAGTCCCTTCATCTCAGCAGCGG TGGAGACTCTCTCTGTGTGTGTCCTGGGCCACTCTACCAGTGATTTCAGACTCCCGCTCTC CCAGCTGTCCTCTCTCATTGTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGG CTCTGGAGTGGGACACTGGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCCGTT GGATCAGACCCTCCTGTGGGCAGGGTTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATA AAAACCAACCCAAATCAA

# FIGURE 304

MFCPLKLILLPVLLDYSLGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLSPG
EHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYICEIRLKGES
QVFKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYY
HKLRMSVEYSQSWGHFQNRVNLVGDIFRNDGSIMLQGVRESDGGNYTCSIHLGNLVFKKTIV
LHVSPEEPRTLVTPAALRPLVLGGNQLVIIVGIVCATILLLPVLILIVKKTCGNKSSVNSTV
LVKNTKKTNPEIKEKPCHFERCEGEKHIYSPIIVREVIEEEEPSEKSEATYMTMHPVWPSLR
SDRNNSLEKKSGGGMPKTQQAF

# FIGURE 305

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATG GTTCTACCCTACTAAAGACAGGAAGATCATAAACTGACAGATACTGAAAATTGTAAGAGTTGG AAACTACATTTTGCAAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCCATCCA GGATGAAGATGGATACATCACCTTAAATATTAAAACTCGGAAACCAGCTCTCGTCTCCGTTG GCCCTGCATCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTGGGG ATGGTTGTCGGGCTGGTGGCTCTGGGGATTTGGTCTGTCATGCAGCGCAATTACCTACAAGA TGAGAATGAAAATCGCACAGGAACTCTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGG TAAAACAATCAGAACTAAAGGGCACTTTCAAAGGTCATAAATGCAGCCCCTGTGACACAAAC TGGAGATATTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACTTAACATGGGAAGAGAG TAAGCAGTACTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGG AGTACATCAAAGCCAGGACTCATTTAATTCGTTGGGTCGGATTATCTCGCCAGAAGTCGAAT GAGGTCTGGAAGTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGGAAGA TGGAAAAGGAAATATGAATTGTGCTTATTTTCATAATGGGAAAATGCACCCTACCTTCTGTG AGAACAAACATTATTTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCT **TAA**TGCAAAGAGGTGGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAAGATATGT ATGAATGCATCAGTAGCTGAAAAAAAAAAAAAAA

# FIGURE 306

MQDEDGYITLNIKTRKPALVSVGPASSSWWRVMALILLILCVGMVVGLVALGIWSVMQRNYL QDENENRTGTLQQLAKRFCQYVVKQSELKGTFKGHKCSPCDTNWRYYGDSCYGFFRHNLTWE ESKQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEFL EDGKGNMNCAYFHNGKMHPTFCENKHYLMCERKAGMTKVDQLP

## FIGURE 307

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGG CGCCGGAGGAGCTCGGACGCCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA GCCCGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCAGCAGCAG CGTCGTGGCCATGCCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCG AGCGCGAGAAATCCAACGCCTGCAAGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGC GACAAAAACAAGTTAAATGTCTTTTCCCGGGTCAAACTCTTCGGCTCCAAGAAGAGGCGCAG AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT CTGTTTAACCTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACTTTTCACACCTGAGTGCA AATTCAAAGAATCAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAG CAGCAGTCAGGCCGAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAA CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC AAGAGCAGAAGTGTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC G**TAG**CCAGTGAGGGCAAAAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAAT CAGAGTTCACTATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

# FIGURE 308

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498

><subunit 1 of 1, 245 aa, 1 stop

><MW: 27564, pI: 10.18, NX(S/T): 1

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLNVFSRVKLFGSKKRRRRP EPQLKGIVTKLYSRQGYHLQLQADGTIDGTKDEDSTYTLFNLIPVGLRVVAIQGVQTKLYLA MNSEGYLYTSELFTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVK KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMSHNEST

### N-glycosylation site.

amino acids 242-246

### Glycosaminoglycan attachment site.

amino acids 165-169, 218-222

### Tyrosine kinase phosphorylation site.

amino acids 93-100

### N-myristoylation site.

amino acids 87-93, 231-237

#### ATP/GTP-binding site motif A (P-loop).

amino acids 231-239

### HBGF/FGF family proteins

amino acids 78-94, 102-153

## FIGURE 309

CCAGGATGGAGCTGGGGCCTGTATAGCCATATTATTGTTCTATGCTACTAGACATGGGGGGG ACTTGGTGAAAAAGGTATTATCCAGCCAGAGGGTCTGGGAGCCCTGTCTTACTGAACCTGGG CAACCTGGATATTCTGAGACATATTTTGGGGGGGATTTCAGTGAAAAAAGTGGGGGATCCCCT CCCCAGTAGGGGTGGGATGAGCGAATATTCCCAAAGCTAAAGTCCCACACCCTGTAGATTAC AAGAGTGGATTTGGCAGGAGTGTGCCCCAAAATACAGTGGAAAGGTGCCTGAAGATATTTAA GAGAGGAGGGAAAGGGGACGTTTTCAATAGGAGGCAAAACTCGAGGGTGGGATCCACTGAGG AGTACATAGGCTGCTGGATCTGGTGGAGCCAGCACTGGGCCCACGGGTGGTAACTGGCTGCT CGAGTCGGGGCCTGAGCGTCAAGAGCATGCCCTAGTGAGCGGGCTCCTCTGGGGGAGCCCAG CGCAGCGGCGCGTGTGTCCCCGCGGCACCAAGTCCCTTTGCCAGAAGCAGCTCCTCATCCTG CTGTCCAAGGTGCGACTGTGCGGGGGGGGGGCCGGCCGGGCCGGACCGGGCCCGGAGCCTCA GCTCAAAGGCATCGTCACCAAACTGTTCTGCCGCCAGGGTTTCTACCTCCAGGCGAATCCCG ACGGAAGCATCCAGGGCACCCCAGAGGATACCAGCTCCTTCACCCACTTCAACCTGATCCCT GTGGGCCTCCGTGTGGTCACCATCCAGAGCGCCAAGCTGGGTCACTACATGGCCATGAATGC TGAGGGACTGCTCTACAGTTCGCCGCATTTCACAGCTGAGTGTCGCTTTAAGGAGTGTGTCT TTGAGAATTACTACGTCCTGTACGCCTCTGCTCTACCGCCAGCGTCGTTCTGGCCGGGCC TGGTACCTCGGCCTGGACAAGGAGGGCCAGGTCATGAAGGGAAACCGAGTTAAGAAGACCAA GGCAGCTGCCCACTTTCTGCCCAAGCTCCTGGAGGTGGCCATGTACCAGGAGCCTTCTCTCC  ${\tt ACAGTGTCCCCGAGGCCTCCCCTTCCAGTCCCCCTGCCCCCT}{\textbf{GA}} {\tt AATGTAGTCCCTGGACTG}$ GAGGTTCCCTGCACTCCCAGTGAGCCAGCCACCACCACCACCTGT

# FIGURE 310

MAALASSLIRQKREVREPGGSRPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDR GPEPQLKGIVTKLFCRQGFYLQANPDGSIQGTPEDTSSFTHFNLIPVGLRVVTIQSAKLGHY MAMNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNR VKKTKAAAHFLPKLLEVAMYQEPSLHSVPEASPSSPPAP

### Tyrosine kinase phosphorylation site:

amino acids 199-207

### N-myristoylation sites:

amino acids 54-60, 89-95, 131-137

### HBGF/FGF family signature:

amino acids 131-155

# FIGURE 311

# FIGURE 312

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503

><subunit 1 of 1, 247 aa, 1 stop

><MW: 27702, pI: 10.36, NX(S/T): 2

MAAAIASGLIRQKRQAREQHWDRPSASRRRSSPSKNRGLCNGNLVDIFSKVRIFGLKKRRLR RQDPQLKGIVTRLYCRQGYYLQMHPDGALDGTKDDSTNSTLFNLIPVGLRVVAIQGVKTGLY IAMNGEGYLYPSELFTPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR VKKTKPAAHFLPKPLEVAMYREPSLHDVGETVPKPGVTPSKSTSASAIMNGGKPVNKSKTT

### N-glycosylation site.

amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 28-32, 29-33

### Tyrosine kinase phosphorylation site.

amino acids 199-207

### N-myristoylation site.

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

### HBGF/FGF family proteins.

amino acids 104-155, 171-198

# FIGURE 313

GAAGGATGCAGGACGCAGCTTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGAAGA ACGAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATG CTCCCCACCCCAAAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAAGTATGTTCATTT TTCTCTATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTTAGTAAAGTAA AGAACTGGTGTGGTGGTGTTTTCCTTTCTTTTTGAATTTCCCACAAGAGGAGAGAAATTAATAATACATCTGC CAGTTGGATTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTT TTTTTAAATTTTTATTCCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTTCTTAACCACCTGGATTTCCATCT GGATGTTGCTGTGATCAGTCTGAAATACAACTGTTTGAATTCCAGAAGGACCAACACCAGATAAATTATGA**ATG** TTGAACAAGATGACCTTACATCCACAGCAGATAATGATAGGTCCTAGGTTTAACAGGGCCCTATTTGACCCCCT GCTTGTGGTGCTGCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGT GCTCCTGCAGCAACCAGTTCAGCAAGGTGATTTGTGTTCGGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCC ACCAACACACGGCTGCTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAG GCACTTGGAAATCCTACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGA ACCTCAACACTCTGGAACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAA CTGAAGGAGCTCTGGTTGCGAAACAACCCCATTGAAAGCATCCCTTCTTATGCTTTTAACAGAATTCCTTCTTT GCGCCGACTAGACTTAGGGGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACT TGAGGTATTTGAACCTTGCCATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAAACTAGATGAG CTGGATCTTTCTGGGAATCATTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGCACCTTCAAAAACT GTGGATGATACAGTCCCAGATTCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCA ACCTGGCACACAATAATCTAACATTACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACAT TTACATCACAACCCTTGGAACTGTAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTC GAACACAGCTTGTTGTGCCCGGTGTAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGA ATTACTTCACATGCTATGCTCCGGTGATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCT GAGCTGAAATGTCGGGCCTCCACATCCCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACA TGGGGCGTACAAAGTGCGGATAGCTGTGCTCAGTGATGGTACGTTAAATTTCACAAATGTAACTGTGCAAGATA CAGGCATGTACACATGTATGGTGAGTAATTCCGTTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCA GCAACCACTACTCCTTTCTCTTACTTTTCAACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACG GACCACAGATAACAATGTGGGTCCCACTCCAGTGGTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACAC CACAGAGCACAAGGTCGACAGAGAAAACCTTCACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATT GATGAGGTCATGAAGACTACCAAAATCATCATTGGGTGTTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCT GGTCATTTCTACAAGATGAGGAAGCAGCACCATCGGCAAAAACCATCACGCCCCAACAAGGACTGTTGAAATTA TTAATGTGGATGATGAGATTACGGGAGACACACCCATGGAAAGCCACCTGCCCATGCCTATCGAGCATGAG CAGTTCAGTGCATGAACCGTTATTGATCCGAATGAACTCTAAAGACAATGTACAAGAGACTCAAATC**TAA**AACA TTTACAGAGTTACAAAAAACAAACAATCAAAAAAAAAAGACAGTTTATTAAAAATGACACAAATGACTGGGCTAA TGATCTAAAGCAGACAAAAA

## FIGURE 314

MLNKMTLHPQQIMIGPRFNRALFDPLLVVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKVIC VRKNLREVPDGISTNTRLLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLDELDLSGNHLSAIRPGSFQGLMHLQKL WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNCDIL WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE LKCRASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVSNSVGN TTASATLNVTAATTTPFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ STRSTEKTFTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYNSYKSPFNHTTTVNTINSIHSS VHEPLLIRMNSKDNVQETQI

#### Signal sequence:

amino acids 1-44

#### Transmembrane domain:

amino acids 523-543

#### N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438, 442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 183-187

### Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

#### N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243, 391-397, 422-428, 433-439, 531-537

# FIGURE 315

GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCCCAGCTCGCCCGAGGTCCGTCGGA GGCGCCCGGCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATC  $\mathsf{GGG}$ CACTGAGATCAAGAGGGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA GTGGTGATCACTTACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG AGTGGCCTTTGCTTCCAATTTCCTGGCAGGAGATGCCTCCTTGCAGATTGAACCTCTGAAGC CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTCAGGGCGCTACGTGTGGAGCCAT GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT ACTGGCAGCGAATCCGAGAGAAAGAGGGAGGAGGATGAACGTCTGCCTCCCAAATCTAGGATT GACTACAACCACCCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA CCAGTGCACAGCAGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACTGTACAGT ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT CCTCTTCCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCCTCCACTCGCTCCACAGCAAAT  $ext{ACGGTC}$ ATTACAATGGACTTGACTCCCACGCTTTCCTAGGAGTCAGGGTCTTTGGACTC TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA GTGAGCATTGCACGGAACAGATTCAGATGAGCATTTTCCTTATACAATACCAAACAAGCAAA AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG AGGTGAATATACCTAAAACTTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTCACAATT TTCAAGAGGAAATGGGATGCTGTTTGTAAATTTTCTATGCATTTCTGCAAACTTATTGGATT ATTAGTTATTCAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC TGAGCTAACCACTTCTAAGAAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC TTCATTTGTCATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAAATAAC TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCCTCAAAT CAGATGCCTCTAAGGACTTTCCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA CCCAACATACCATTATAGTCTCTTCTTTCTGAGAAAATGTGAAAACCAGAATTGCAAGACTGG TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

# FIGURE 316

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV ILKVLVRPSKPKCELEGELTEGSDLTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI FLLVWLLIRRKDKERYEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

### Signal sequence:

amino acids 1-16

#### Transmembrane domain:

amino acids 232-251

# FIGURE 317

 $\texttt{CC} \underline{\textbf{ATG}} \texttt{GCGCTCCTGCTGCTGCTGCTGCGGAGTAGTGGATTTCGCCAGAAGTTTGAGTATCACTACT}$ CCTGAAGAGATGATTGAAAAAGCCAAAGGGGAAACTGCCTATCTGCCATGCAAATTTACGCTTAGTCCCGAAGA ATTCTGGAGACAAAATTTATGATGACTACTATCCAGATCTGAAAGGCCGAGTACATTTTACGAGTAATGATCTC AAATCTGGTGATGCATCAATAAATGTAACGAATTTACAACTGTCAGATATTGGCACATATCAGTGCAAAGTGAA AAAAGCTCCTGGTGTTGCAAATAAGAAGATTCATCTGGTAGTTCTTGTTAAGCCTTCAGGTGCGAGATGTTACG TTGATGGATCTGAAGAAATTGGAAGTGACTTTAAGATAAAATGTGAACCAAAAGAAGGTTCACTTCCATTACAG TATGAGTGGCAAAAATTGTCTGACTCACAGAAAATGCCCACTTCATGGTTAGCAGAAATGACTTCATCTGTTAT ATCTGTAAAAATGCCTCTTCTGAGTACTCTGGGACATACAGCTGTACAGTCAGAAACAGAGTGGGCTCTGATC AGTGCCTGTTGCGTCTAAACGTTGTCCCTCCTTCAAATAAAGCTGGACTAATTGCAGGAGCCATTATAGGAACT GGAAGTTCATCACGATATCAGGGAAGATGTGCCACCTCCAAAGAGCCGTACGTCCACTGCCAGAAGCTACATCG GCAGTAATCATTCATCCCTGGGGTCCATGTCTCCTTCCAACATGGAAGGATATTCCAAGACTCAGTATAACCAA GTACCAAGTGAAGACTTTGAACGCACTCCTCAGAGTCCGACTCTCCCACCTGCTAAGTTCAAGTACCCTTACAA  $\texttt{GACTGATGGAATTACAGTTGTA} \underline{\textbf{TAA}} \texttt{ATATGGACTACTGAAGAATCTGAAGTATTGTATTATTTGACTTTATTTT}$ AGGCCTCTAGTAAAGACTTAAATGTTTTTTAAAAAAAGCACAAGGCACAGAGATTAGAGCAGCTGTAAGAACAC ATCTACTTTATGCAATGGCATTAGACATGTAAGTCAGATGTCATGTCAAAATTAGTACGAGCCAAATTCTTTGT TAAAAAACCCTATGTATAGTGACACTGATAGTTAAAAGATGTTTTATTATATTTTCAATAACTACCACTAACAA ATTTTTAACTTTTCATATGCATATTCTGATATGTGGTCTTTTAGGAAAAGTATGGTTAATAGTTGATTTTTCAA AGGAAATTTTAAAATTCTTACGTTCTGTTTAATGTTTTTGCTATTTAGTTAAATACATTGAAGGGAAATACCCG TAAAGTAAATTCTCAAAGGTGCTAGAACAAATCGTCCACTTCTACAGTGTTCTCGTATCCAACAGAGTTGATGC ACAATATATAAATACTCAAGTCCAATATTAAAAACTTAGGCACTTGACTAACTTTAATAAAATTTCTCAAACTA TATCAATATCTAAAGTGCATATATTTTTTAAGAAAGATTATTCTCAATAACTTCTATAAAAATAAGTTTGATGG TTTGGCCCATCTAACTTCACTACTATTAGTAAGAACTTTTAACTTTTAATGTGTAGTAAGGTTTATTCTACCTT TTTCTCAACATGACACCAACACAATCAAAAACGAAGTTAGTGAGGTGCTAACATGTGAGGATTAATCCAGTGAT TCCGGTCACAATGCATTCCAGGAGGAGGTACCCATGTCACTGGAATTGGGCGATATGGTTTATTTTTTCTTCCC TGATTTGGATAACCAAATGGAACAGGAGGAGGATAGTGATTCTGATGGCCATTCCCTCGATACATTCCTGGCTT TTTTCTGGGCAAAGGGTGCCACATTGGAAGAGGTGGAAATATAAGTTCTGAAATCTGTAGGGAAGAGAACACAT TAAGTTAATTCAAAGGAAAAAATCATCATCTATGTTCCAGATTTCTCATTAAAGACAAAGTTACCCACAACACT GAGATCACATCTAAGTGACACTCCTATTGTCAGGTCTAAATACATTAAAAACCTCATGTGTAATAGGCGTATAA TGTATAACAGGTGACCAATGTTTTCTGAATGCATAAAGAAATGAATAAACTCAAACACAGTACTTCCTAAACAA CTTCAACCAAAAAAGACCAAAACATGGAACGAATGGAAGCTTGTAAGGACATGCTTGTTTTAGTCCAGTGGTTT CCACAGCTGGCTAAGCCAGGAGTCACTTGGAGGCTTTTAAATACAAAACATTGGAGCTGGAGGCCATTATCCTT AGCAAACTAATGCAGAAACAGAAAATCAACTACCGCATGTTCTCACTTATAAGTGGGAGGTAATGATAAGAACT TATGAACACAAAGAAGAACAATAGACATTGGAGTCTATTTGAGAGGGGAGGGTGGGAGAAGGAAAAGGAGCA GAAAAGATAACTATTGAGTACTGCCTTCACACCTGGGTGATGAAATAATATGTACAACAAATCCCTGTGACACA 

## FIGURE 318

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA82361

><subunit 1 of 1, 352 aa, 1 stop

><MW: 38938, pI: 7.86, NX(S/T): 3

MALLLCFVLLCGVVDFARSLSITTPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLISPA DNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK APGVANKKIHLVVLVKPSGARCYVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSDSQKMPT SWLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSDQCLLRLNVVPPSNKAGLIAGAIIGTLL ALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHSSLGSMSPSNM EGYSKTQYNQVPSEDFERTPQSPTLPPAKFKYPYKTDGITVV

#### Signal sequence.

amino acids 1-19

#### Transmembrane domain:

amino acids 236-257

### N-glycosylation sites.

amino acids 106-110, 201-205, 298-302

### Tyrosine kinase phosphorylation sites.

amino acids 31-39, 78-85, 262-270

#### N-myristoylation sites.

amino acids 116-122, 208-214, 219-225, 237-243, 241-247, 245-251, 296-302

### Myelin P0 protein.

amino acids 96-125

# FIGURE 319

CTCAAGCATCACTTACAGGACCAGAGGGACAAGACATGACTGTGATGAGGAGCTGCTTTCGC CAATTTAACACCAAGAAGAATTGAGGCTGCTTGGGAGGAAGGCCAGGAGGAACACGAGACTG AGAG**ATG**AATTTTCAACAGAGGCTGCAAAGCCTGTGGACTTTAGCCAGACCCTTCTGCCCTC CTTTGCTGGCGACAGCCTCTCAAATGCAGATGGTTGTGCTCCCTTGCCTGGGTTTTACCCTG CTTCTCTGGAGCCAGGTATCAGGGGCCCAGGGCCAAGAATTCCACTTTGGGCCCTGCCAAGT GAAGGGGGTTGTTCCCCAGAAACTGTGGGAAGCCTTCTGGGCTGTGAAAGACACTATGCAAG CTCAGGATAACATCACGAGTGCCCGGCTGCTGCAGCAGGAGGTTCTGCAGAACGTCTCGGAT GCTGAGAGCTGTTACCTTGTCCACACCCTGCTGGAGTTCTACTTGAAAACTGTTTTCAAAAA CCACCACAATAGAACAGTTGAAGTCAGGACTCTGAAGTCATTCTCTACTCTGGCCAACAACT TTGTTCTCATCGTGTCACAACTGCAACCCAGTCAAGAAAATGAGATGTTTTCCATCAGAGAC AGTGCACACAGGCGGTTTCTGCTATTCCGGAGAGCATTCAAACAGTTGGACGTAGAAGCAGC TCTGACCAAAGCCCTTGGGGAAGTGGACATTCTTCTGACCTGGATGCAGAAATTCTACAAGC  $\verb"TCTGA" ATGTCTAGACCAGGACCTCCCTCCCCCTGGCACTGGTTTGTTCCCTGTGTCATTTCA"$ AACAGTCTCCCTTCCTATGCTGTTCACTGGACACTTCACGCCCTTGGCCATGGGTCCCATTC TTGGCCCAGGATTATTGTCAAAGAAGTCATTCTTTAAGCAGCGCCAGTGACAGTCAGGGAAG AATTAATGTCAGTATTTCAACTGAAGTTCTATTTATTTGTGAGACTGTAAGTTACATGAAGG CAGCAGAATATTGTGCCCCATGCTTCTTTACCCCTCACAATCCTTGCCACAGTGTGGGGCAG GTTAAAAAACAGAGAGGGATGCTTGGATGTAAAACTGAACTTCAGAGCATGAAAATCACACT GTCTTCTGATATCTGCAGGGACAGAGCATTGGGGGTGGGGGTAAGGTGCATCTGTTTGAAAAG TAAACGATAAAATGTGGATTAAAGTGCCCAGCACAAAGCAGATCCTCAATAAACATTTCATT TATCCTAGTCATTCTTCCCTAATCTTCCACTTGAGTGTCAAGCTGACCTTGCTGATGGTGAC ATTGCACCTGGATGTACTATCCAATCTGTGATGACATTCCCTGCTAATAAAAGACAACATAA СТССААААААААААААААААААААААААА

# FIGURE 320

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002

><subunit 1 of 1, 206 aa, 1 stop

><MW: 23799, pI: 9.12, NX(S/T): 3

MNFQQRLQSLWTLARPFCPPLLATASQMQMVVLPCLGFTLLLWSQVSGAQGQEFHFGPCQVK GVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLVHTLLEFYLKTVFKNH HNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRFLLFRRAFKQLDVEAAL TKALGEVDILLTWMQKFYKL

### Signal sequence:

amino acids 1-42

### N-glycosylation sites.

amino acids 85-89, 99-103, 126-130

# FIGURE 321

# FIGURE 322

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282

><subunit 1 of 1, 177 aa, 1 stop

><MW: 20452, pI: 8.00, NX(S/T): 2

 $\label{thm:local} \\ \mathbf{MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST} \\ \mathbf{LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQ} \\ \mathbf{RQCHCRQEATNATRVIHDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA}$ 

### Signal sequence:

amino acids 1-18

### N-glycosylation sites.

amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

### N-myristoylation site.

amino acids 24-30

Actinin-type actin-binding domain signature 1.

amino acids 159-169

# FIGURE 323

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAG AACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTAT AGAATAACATCCACTTTGCCTTTCTCCCACAGGTGTCCACTCCCAGGTCCAACTGCACCTC GGTTCTATCGATAATCTCAGCACCAGCCACTCAGAGCAGGGCACGATGTTGGGGGGCCCGCCT CAGGCTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCA ATGCCTCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCC AGGAACAGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGAC CATCTACAGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGA TGAGCAGAAGATACCTCTGCATGGATTTCAGAGGCAACATTTTTGGATCACACTATTTCGAC CCGGAGAACTGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCC TCAGTATCACTTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACC CACCCCGTACTCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACTTCAACACC CCCATACCACGGCGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCCTGAACGT GCTGAAGCCCCGGGCCCGGATGACCCCGGCCCCGGCCTCCTGTTCACAGGAGCTCCCGAGCG CCGAGGACAACAGCCCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTCGAGTGAAC  ${\tt ACGCACGCTGGGGGAACGGCCCGGAAGGCTGCCGCCCCTTCGCCAAGTTCATC} {\bf \underline{TAG}} {\tt GGTCG}$ CTGG

# FIGURE 324

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142238</pre>

><subunit 1 of 1, 251 aa, 1 stop

><MW: 27954, pI: 9.22, NX(S/T): 1

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSSWGGLIHLYTATARNSYHLQIHKNGHVD GAPHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLENGY DVYHSPQYHFLVSLGRAKRAFLPGMNPPPYSQFLSRRNEIPLIHFNTPIPRRHTRSAEDDSE RDPLNVLKPRARMTPAPASCSQELPSAEDNSPMASDPLGVVRGGRVNTHAGGTGPEGCRPFA KFI

### Important features of the protein:

#### Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 175-179

#### N-myristoylation site.

amino acids 33-39, 100-106, 225-231, 229-235

#### HBGF/FGF family proteins

amino acids 73-124

## FIGURE 325

GGAAAAGGTACCCGCGAGAGACAGCCAGCAGTTCTGTGGAGCAGCGGTGGCCGGCTAGGATG GAGCTCTGCAGGCCCCAGCACCCGCAGAGCAGACACTGCGATGACAACGGACGACACAGAAG TGCCCGCTATGACTCTAGCACCGGGCCACGCCGCTCTGGAAACTCAAACGCTGAGCGCTGAG ACCTCTTCTAGGGCCTCAACCCCAGCCGGCCCCATTCCAGAAGCAGAGCCAGGGGAGCCAA GAGAATTTCCCCTGCAAGAGAGACCAGGAGTTTCACAAAAACATCTCCCAACTTCATGGTGC TGATCGCCACCTCCGTGGAGACATCAGCCGCCAGTGGCAGCCCCGAGGGAGCTGGAATGACC ACAGTTCAGACCATCACAGGCAGTGATCCCGAGGAAGCCATCTTTGACACCCTTTGCACCGA TGACAGCTCTGAAGAGGCAAAGACACTCACAATGGACATATTGACATTGGCTCACACCTCCA CAGAAGCTAAGGGCCTGTCCTCAGAGAGCAGTGCCTCTTCCGACGGCCCCCATCCAGTCATC ACCCCGTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCC GTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCCGTCAT GGTCCCCGGGATCTGATGTCACTCTCCTCGCTGAAGCCCTGGTGACTGTCACAAACATCGAG GTTATTAATTGCAGCATCACAGAAATAGAAACAACAACTTCCAGCATCCCTGGGGCCTCAGA CATAGATCTCATCCCCACGGAAGGGGTGAAGGCCTCGTCCACCTCCGATCCACCAGCTCTGC CTGACTCCACTGAAGCAAAACCACACATCACTGAGGTCACAGCCTCTGCCGAGACCCTGTCC ACAGCCGGCACCACAGAGTCAGCTGCACCTCATGCCACGGTTGGGACCCCACTCCCCACTAA CAGCGCCACAGAAGAGAGTGACAGCACCCGGGGCCACGACCCTCAGTGGAGCTCTGGTCA CAGTTAGCAGGAATCCCCTGGAAGAAACCTCAGCCCTCTCTGTTGAGACACCAAGTTACGTC AAAGTCTCAGGAGCAGCTCCGGTCTCCATAGAGGCTGGGTCAGCAGTGGGCAAAACAACTTC CTTTGCTGGGAGCTCTGCTTCCTCCTACAGCCCCTCGGAAGCCGCCCTCAAGAACTTCACCC CTTCAGAGACACCGACCATGGACATCGCAACCAAGGGGCCCTTCCCCACCAGCAGGGACCCT CTTCCTTCTGTCCCTCCGACTACAACCAACAGCAGCCGAGGGACGAACAGCACCTTAGCCAA GATCACAACCTCAGCGAAGACCACGATGAAGCCCCAACAGCCACGCCCACGACTGCCCGGAC  ${\tt GAGGCCGACCACAGACG}{\textbf{TGA}}{\tt GTGCAGGTGAAAATGGAGGTTTCCTCCTCCTGCGGCTGAGTG}$ TGGCTTCCCCGGAAGACCTCACTGACCCCAGAGTGGCAGAAAGGCTGATGCAGCAGCTCCAC CGGGAACTCCACGCCCACGCGCCTCACTTCCAGGTCTCCTTACTGCGTGTCAGGAGAGGCTA ACGGACATCAGCTGCAGCCAGGCATGTCCCGTATGCCAAAAGAGGGTGCTGCCCCTAGCCTG GGCCCCACCGACAGACTGCAGCTGCGTTACTGTGCTGAGAGGTACCCAGAAGGTTCCCATG AAGGGCAGCATGTCCAAGCCCCTAACCCCAGATGTGGCAACAGGACCCTCGCTCACATCCAC  $\tt CGGAGTGTATGTATGGGGGGGGGTTCACCTGTTCCCAGAGGTGTCCTTGGACTCACCTTGG$ CACATGTTCTGTGTTTCAGTAAAGAGAGACCTGATCACCCATCTGTGTGCTTCCATCCTGCA TTAAAATTCACTCAGTGTGGCCCAAAAAAAA

## FIGURE 326

MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTLSA
ETSSRASTPAGPIPEAETRGAKRISPARETRSFTKTSPNFMVLIATSVETSAASGSPEGAGM
TTVQTITGSDPEEAIFDTLCTDDSSEEAKTLTMDILTLAHTSTEAKGLSSESSASSDGPHPV
ITPSRASESSASSDGPHPVITPSRASESSASSDGPHPVITPSWSPGSDVTLLAEALVTVTNI
EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL
STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTLSGALVTVSRNPLEETSALSVETPSY
VKVSGAAPVSIEAGSAVGKTTSFAGSSASSYSPSEAALKNFTPSETPTMDIATKGPFPTSRD
PLPSVPPTTTNSSRGTNSTLAKITTSAKTTMKPQQPRPRLPGRGRPQT

### N-glycosylation sites:

amino acids 252-256, 445-449, 451-455

cAMP-and cGMP-dependent protein kinase phosphorylation site. amino acids 84-90

### Casein kinase II phosphorylation sites.

amino acids 37-41, 108-112, 131-135, 133-137, 148-152, 165-169, 246-250, 254-258, 256-260, 269-273, 283-287, 333-337, 335-339, 404-408, 414-418, 431-435

#### N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284, 314-320, 349-355, 386-392, 397-403, 449-455

#### ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

## FIGURE 327

GCGGAGCATCCGCTGCGGTCCTCGCCGAGACCCCCGCGCGGATTCGCCGGTCCTTCCCGCGG GCGCGACAGAGCTGTCCTCGCACCTGGATGGCAGCAGGGGGCGCCGGGGTCCTCTCGACGCCA CCTTGACCTTTGAAGACCAAAACTAAACTGAAATTTAAA**ATG**TTCTTCGGGGGGAGAAGGGAG CTTGACTTACACTTTGGTAATAATTTGCTTCCTGACACTAAGGCTGTCTGCTAGTCAGAATT GCCTCAAAAAGAGTCTAGAAGATGTTGTCATTGACATCCAGTCATCTCTTTCTAAGGGAATC AGAGGCAATGAGCCCGTATATACTTCAACTCAAGAAGACTGCATTAATTCTTGCTGTTCAAC AAAAACATATCAGGGGACAAAGCATGTAACTTGATGATCTTCGACACTCGAAAAACAGCTA GACAACCCAACTGCTACCTATTTTTCTGTCCCAACGAGGAAGCCTGTCCATTGAAACCAGCA AAAGGACTTATGAGTTACAGGATAATTACAGATTTTCCATCTTTGACCAGAAATTTGCCAAG CCAAGAGTTACCCCAGGAAGATTCTCTCTTACATGGCCAATTTTCACAAGCAGTCACTCCCC TAGCCCATCATCACACAGATTATTCAAAGCCCACCGATATCTCATGGAGAGACACACTTTCT GCTCCTTGCTTATAAGGAAAAAGGCCATTCTCAGAGTTCACAATTTTCCTCTGATCAAGAAA TAGCTCATCTGCTGCCTGAAAATGTGAGTGCGCTCCCAGCTACGGTGGCAGTTGCTTCTCCA CATACCACCTCGGCTACTCCAAAGCCCGCCACCCTTCTACCCACCAATGCTTCAGTGACACC TTCTGGGACTTCCCAGCCACAGCTGGCCACCACAGCTCCACCTGTAACCACTGTCACTTCTC AGCCTCCACGACCCTCATTTCTACAGTTTTTTACACGGGCTGCGGCTACACTCCAAGCAATG GCTACAACAGCAGTTCTGACTACCACCTTTCAGGCACCTACGGACTCGAAAGGCAGCTTAGA AACCATACCGTTTACAGAAATCTCCAACTTAACTTTGAACACAGGGAATGTGTATAACCCTA CTGCACTTTCTATGTCAAATGTGGAGTCTTCCACTATGAATAAAACTGCTTCCTGGGAAGGT AGGGAGGCCAGTCCAGGCAGTTCCTCCCAGGGCAGTGTTCCAGAAAATCAGTACGGCCTTCC ATTTGAAAAATGGCTTCTTATCGGGTCCCTGCTCTTTTGGTGTCCTGTTCCTGGTGATAGGCC TCGTCCTCCTGGGTAGAATCCTTTCGGAATCACTCCGCAGGAAACGTTACTCAAGACTGGAT TATTTGATCAATGGGATCTATGTGGACATC**TAA**GGATGGAACTCGGTGTCTCTTAATTCATT TAGTAACCAGAAGCCCAAATGCAATGAGTTTCTGCTGACTTGCTAGTCTTAGCAGGAGGTTG GCTCTGTTGCCCAGGCTGGAGTGCAGTAGCACGATCTCGGCTCTCACCGCAACCTCCGTCTC CTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTAAGTATCTGGGATTACAGGCATGTGCCA CCACACCTGGGTGATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTG GTCTCAAACTCCTGACCTAGTGATCCACCCTCCTCGGCCTCCCAAAGTGCTGGGATTACAGG CATGAGCCACCACAGCTGGCCCCCTTCTGTTTTATGTTTGGTTTTTGAGAAGGAATGAAGTG GGAACCAAATTAGGTAATTTTGGGTAATCTGTCTCTAAAATATTAGCTAAAAAACAAAGCTCT ATGTAAAGTAATAAAGTATAATTGCCATATAAATTTCAAAATTCAACTGGCTTTTATGCAAA GAAACAGGTTAGGACATCTAGGTTCCAATTCATTCACATTCTTGGTTCCAGATAAAATCAAC TGTTTATATCAATTTCTAATGGATTTGCTTTTTTTTATATGGATTCCTTTAAAACTTATT CCAGATGTAGTTCCTTCCAATTAAATATTTGAATAAATCTTTTGTTACTCAA

## FIGURE 328

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410

><subunit 1 of 1, 431 aa, 1 stop

><MW: 46810, pI: 6.45, NX(S/T): 6

MFFGGEGSLTYTLVIICFLTLRLSASQNCLKKSLEDVVIDIQSSLSKGIRGNEPVYTSTQED
CINSCCSTKNISGDKACNLMIFDTRKTARQPNCYLFFCPNEEACPLKPAKGLMSYRIITDFP
SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF
KMDEASAQLLAYKEKGHSQSSQFSSDQEIAHLLPENVSALPATVAVASPHTTSATPKPATLL
PTNASVTPSGTSQPQLATTAPPVTTVTSQPPTTLISTVFTRAAATLQAMATTAVLTTTFQAP
TDSKGSLETIPFTEISNLTLNTGNVYNPTALSMSNVESSTMNKTASWEGREASPGSSSQGSV
PENQYGLPFEKWLLIGSLLFGVLFLVIGLVLLGRILSESLRRKRYSRLDYLINGIYVDI

### Signal sequence.

amino acids 1-25

### Transmembrane domain.

amino acids 384-405

### N-glycosylation sites.

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 415-419

### Tyrosine kinase phosphorylation site.

amino acids 50-57

### N-myristoylation sites.

amino acids 4-10, 48-54, 315-321

# FIGURE 329

CCCAGGTTATGAAGCCCTGGAGGGCCCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT GGGATCCTCTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGACCCAGGAGACAAT GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA ACCTCACCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG TCTTTACTGATCTCTGTTCGTCTTTCCAGGACCCTGCTGTCCCCCTCCCCTTCTCCCAC CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCC CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACCAGCCAAGCAGGGGAAGACAGGG GCTGAGGCCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACAC AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCATGCAGC TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCCTGGTGCTGAGCCTTCTGTCAGC CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC CCTTCCCAGGCCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGA  $\texttt{GCTGGGCTTCTCGAAGTTTGTCTCAGCG} \underline{\textbf{TAG}} \texttt{GGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT}$ GAAGCAGTATGGCTGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAG TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCAGGCTCTCCTCTTGCATGTTCCA GCCTGACCTAGAAGCGTTTGTCAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG GAGACTGGGACATCCCTGATAGGTTCACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACTCCTGGGC TGGCGTCCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT GTGAAAAACGTGATTCCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG GACTCTGAATTCTAACAATGCCCAGTGACTGTCGCACTTGAGTTTGAGGGCCAGTGGGCCTG ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG TCCAGGCCTTGGTCAGGTCAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCTTTGGAAAAAATGATGAAGA AAACCTTGGCTCCTTCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTTCTGGTGGCTAGAGA GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTCGGGGGTGGTAAAGTA GCACAACTACTATTTTTTTTTTTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACTCCGCCTCCTGGGTTCAAGTGATT TTTGTACTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGAC CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG TCTGGCCCTATTTCCTTTAAAAAGTGAAATTAAGAGTTGTTCAGTATGCAAAACTTGGAAAG ATGGAGGAGAAAAAGGAAAAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT TATTTCGTTTTGTTGTACTTCCTTCCACTCTTTTCTTCTTCACATAATTTGCCGGTGTTCTT TTTACAGAGCAATTATCTTGTATATACAACTTTGTATCCTGCCTTTTCCACCTTATCGTTCC GCTGCATAAAAAAAAAAAAAAAA

# FIGURE 330

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196</pre>

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS GTIYAEEEGQETMKGRVSIRDSRQELSLIVTLWNLTLQDAGEYWCGVEKRGPDESLLISLFV FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG TSQYGHERTSQYTGTSPHPATSPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAEEKEAPSQAPEGD VISMPPLHTSEEELGFSKFVSA

#### Important features:

#### Signal peptide:

amino acids 1-17

### Transmembrane domain:

amino acids 248-269

### N-glycosylation site.

amino acids 96-99

# Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

### Ig like V-type domain:

amino acids 13-128